
Electronic Thesis and Dissertation Repository

8-22-2014 12:00 AM

Characterizing Spoken Discourse in Individuals with Parkinson Disease Without Dementia

Angela C. Roberts
The University of Western Ontario

Supervisor
JB Orange
The University of Western Ontario

Graduate Program in Health and Rehabilitation Sciences
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy
© Angela C. Roberts 2014

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Speech Pathology and Audiology Commons](#)

Recommended Citation

Roberts, Angela C., "Characterizing Spoken Discourse in Individuals with Parkinson Disease Without Dementia" (2014). *Electronic Thesis and Dissertation Repository*. 2336.
<https://ir.lib.uwo.ca/etd/2336>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Charcterizing Spoken Discourse in Individuals with Parkinson Disease Without
Dementia

(Thesis format: Monograph)

by

Angela Christine Roberts (South)

Graduate Program in Health and Rehabilitation Sciences

A thesis submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Angela Roberts, 2014

Abstract

Background: The effects of disease (PD) on cognition, word retrieval, syntax, and speech/voice processes may interact to manifest uniquely in spoken language tasks. A handful of studies have explored spoken discourse production in PD and, while not ubiquitously, have reported a number of impairments including: reduced words per minute, reduced grammatical complexity, reduced informativeness, and increased verbal disruption. Methodological differences have impeded cross-study comparisons. As such, the profile of spoken language impairments in PD remains ambiguous.

Method: A cross-genre, multi-level discourse analysis, prospective, cross-sectional between groups study design was conducted with 19 PD participants ($M_{\text{age}} = 70.74$, $M_{\text{UPDRS-III}} = 30.26$) and 19 healthy controls ($M_{\text{age}} = 68.16$) without dementia. The extensive protocol included a battery of cognitive, language, and speech measures in addition to four discourse tasks. Two tasks each from two discourse genres (picture sequence description; story retelling) were collected. Discourse samples were analysed using both microlinguistic and macrostructural measures. Discourse variables were collapsed statistically to a primal set of variables used to distinguish the spoken discourse of PD vs. controls.

Results: Participants with PD differed significantly from controls along a continuum of productivity, grammar, informativeness, and verbal disruption domains including total words $F(1,36) = 3.87, p = .06$; words/minute $F(1,36) = 7.74, p = .01$, % grammatical utterances $F(1,36) = 11.92, p = .001$, total CIUs $F(1,36) = 13.30, p = .001$, % CIUs (Correct Information Units) $F(1,36) = 9.35, p = .004$, CIUs/minute $F(1,36) = 14.06, p = .001$, and verbal disruptions/100 words $F(1,36) = 3.87, p = .06$ ($\alpha = .10$). Discriminant function analyses showed that optimally weighted discourse variables discriminated the spoken discourse of PD vs. controls with 81.6% sensitivity and 86.8% specificity. For both discourse genres, discourse performance showed robust, positive, correlations with global cognition. In PD (picture sequence description), more impaired discourse performance correlated significantly with more severe motor impairment, more advanced disease staging, and higher doses of PD medications.

Conclusions: The spoken discourse in PD without dementia differs significantly and predictably from controls. Results have both research and clinical implications.

Keywords: Parkinson disease, movement disorders, spoken discourse, language, picture description, story retelling, spontaneous language, cognition, multivariate analysis, discriminant function analysis

Epigraph

“Language is the blood of the soul into which thoughts run and out of which they grow.”

Dr. Oliver Wendell Holmes (1809 – 1894)

“Words mean more than what is set down on paper. It takes the human voice to infuse them with deeper meaning.”

Dr. Maya Angelou (1928 – 2014)

Dedication

This thesis is dedicated to those individuals with Parkinson disease and their care partners who do not *yield*, but who instead face each day with courage and choose to learn to *live* and *conquer* this disease. Thank you for inspiring me with your stories.

Just like moons and like suns,
With the certainty of tides,
Just like hopes springing high,
Still I'll rise.

From: *Still I'll Rise* by Dr. Maya Angelou

Acknowledgements

Dr. JB Orange, for his incredible personal and academic support. I will be forever grateful to him for his generosity, his kindness, his mentorship, and his integrity. He is a true gentleman, scholar, steward of the academic process, and trusted friend.

Dr. Mandar Jog who provided the courage and inspiration to pursue this path. Thank you for your mentorship, guidance, and friendship.

Dr. Scott Adams who has served as a valuable source of guidance, focus, and confidence throughout my PhD process.

Professor Emeritus, Dr. Marilyn Kertoy, who is a valuable mentor and friend. She is and has been a beacon of support throughout this process.

Dr. Andrew Johnson for his guidance in designing the plan for statistical analyses.

Dr. Sheila Moodie for her guidance in the incorporation of principles and theories of knowledge translation into my public lecture.

The funding sources for this project included the Parkinson Society Canada Graduate Student Award and The Canadian Institutes of Health Research Fellowship Award and Prize in Aging. Without their financial support this process would not have been possible.

Parkinson Society Southwest Ontario for giving me a ‘voice’ for so many years and believing in the message that I had to share.

Katie Findlater, Bracia Eaton, and Shalane Basque for their assistance with the transcription component of the study. They met every task with enthusiasm.

Victoria Mains, Megan Hansen, and Tiffany Ashford for their incredible support and assistance with the reliability components of the study.

The numerous audiology and speech-language pathology students who participated in the study as scorers and raters.

The Aging and Communication Laboratory at Western University, Dr. JB Orange.

The Movement Disorders Laboratory at London Health Sciences Centre, Dr. Mandar Jog.

The faculty and staff of Communication Sciences and Disorders. They have been a well of never-ending support through challenging times.

The faculty and staff of Health and Rehabilitation Sciences for guidance in navigating the PhD process.

Dr. Jack Scott and Dr. JB Orange for assistance with the editing process.

Last, but never least, my incredibly supportive friends and family. While this list is not exhaustive, a few special mentions are warranted:

To Jack, who never allowed me give up and always believed in me. You have been a beacon of light in the dark of night and provided rest for my weary head.

To my parents, who helped me take that first step, and then always kept me facing forward.

To Hillary, my Anam Cara, my sister, my friend. I could not have done this without you.

To the Brown-Dalton Clan, my extended Canadian family and my support - especially to Tom and Wendy who have given me shelter in times of need and have always been there to keep me focused on positive days ahead.

To those who have been voices of support and welcoming ears in times of need. Your friendship is invaluable: Stephanie Somers, Mikelle Meaden, Dr. Stella Ng, Dr. Jeff Crukley, Dr. Susan Scollie, Dr. Sheila Moodie, Rebecca Affoo.

Table of Contents

Abstract	ii
Epigraph	iv
Dedication.....	v
Acknowledgements	vi
List of Tables	x
List of Figures	xi
List of Appendices	xii
List of Abbreviations	xiii
Chapter 1: Thesis Overview	1
1.1 Study Objective.....	2
1.2 Organization of Thesis Chapters	2
Chapter 2: Background.....	4
2.1 Parkinson Disease: An Overview.....	4
2.2 Cognition, Language, and Motor Speech Impairments	9
2.3 Studies of Spoken Monologic Discourse in PD	16
2.4 Background Summary.....	47
2.5 Rationale for the Current Study and Statement of the Problem	48
2.6 Research Questions (RQ), Hypotheses, and Objectives	48
Chapter 3: Method.....	50
3.1 Ethics Approval, Participant Recruitment and Enrolment.....	50
3.2 Description of Participants	54
3.3 Audio Recording Protocol.....	57
3.4 Data Collection Environment.....	58
3.5 General Method	59
3.6 Hearing Screening	62
3.7 Vision Screening.....	62
3.8 Collection Demographic Information	63
3.9 Collection of Disease Related Measures.....	63
3.10 Measures of Cognition and Language.....	64
3.11 Measures of Speech Intelligibility and Rate	68
3.12 Elicitation of Discourse Samples.....	71
3.13 Transcription, Coding and Segmenting of Discourse Samples	76
3.14 Discourse Analyses	79
3.15 Statistical Analyses	84
Chapter 4: Results	92
4.1 Cognition, Language and Speech Intelligibility Measures (RQ 1).....	92
4.2 Discourse Measures in Relation to Group x Task and Task Effects (RQ 2)	95
4.3 Discourse Measures in Relation to Group (RQ 3)	104
4.3 Discourse Discriminant Function (RQ 4).....	109

4.4 Relationship Between Discourse Performance and Disease Variables, Age, Education and Cognition (RQ 5)	115
Chapter 5: Discussion	117
5.1 Participant Sampling and Description Considerations	117
5.2 Discourse Sampling Considerations.....	127
5.3 Discussion of Findings in Relation to Specific Research Questions.....	130
5.4 Study Limitations.....	161
5.5 Theoretical and Clinical Relevance of the Findings	164
5.6 Future Directions.....	166
5.7 Conclusion.....	168
References	169

List of Tables

Table Title	Page(s)
Table 1. Diagnostic Features of Idiopathic Parkinson Disease (PD)	5
Table 2. Summary of Published Discourse Production Studies in Parkinson Disease	19-22
Table 3. Study Inclusion and Exclusion Criteria	53
Table 4. Demographic Data for Participants by Group	55
Table 5. Overview of Study Tasks and Order of Administration	60-61
Table 6. Standardized Measures of Cognition and Language Administered in the Study Protocol	64-65
Table 7. Multi-level Discourse Analyses Protocol	81-82
Table 8. Inter-rater and Intra-rater ICC Results for Discourse Measure Reliability Study	83-84
Table 9. Performance on Cognitive and Language Measures for participants with PD and Controls (Ctrl)	93-94
Table 10. Summary of Levene's Tests for all Discourse Variables	97-98
Table 11. Discourse Performance Effects in Relation to Task Independent of Group	99-101
Table 12. Discourse Performance In Relation to Group Independent of Task	104-105
Table 13. Summary of the Assessed Discriminant Functions	112
Table 14. Summary of Overall and Cross-Validated Predictive Accuracies for PD Discourse	113-114

List of Figures

Figure Title	Page(s)
<i>Figure 1.</i> Picture sequence description 1: ‘Argument’	72
<i>Figure 2.</i> Picture sequence description 2: ‘Directions’	73
<i>Figure 3.</i> Transcript of story retelling 1: ‘The Glass of Water’	74
<i>Figure 4.</i> Transcript of story retelling 2: ‘Out of Gas’	75
<i>Figure 5.</i> Summary of variance associated with each statistically significant canonical variable in relation to Task	101
<i>Figure 6.</i> Summary of proportion of variance associated with Group for statistically significant canonical variable	106
<i>Figure 7.</i> Classification of individual discourse samples using the Discourse Discriminant Function	114

List of Appendices

Appendix Title	Page(s)
Appendix A. Western University Ethics Approval	187
Appendix B. Lawson Health Research Institutes Approval	188
Appendix C. Complete discourse protocol from which the tasks in the current study are abstracted	189-190
Appendix D. CHAT Symbols Used in Coding Discourse Transcripts	191-192
Appendix E. Conventions for Segmenting C-units	293-195
Appendix F. Sample of an Orthographically Transcribed, Coded, and Segmented Audio File Prepared for Discourse Analyses	196
Appendix G. Discourse Analysis Form	197
Appendix H. Discourse Analyses Definitions/Procedures	198-203
Appendix I. Rules for Calculating CIUs	204-206

List of Abbreviations

ABCD-VL	Arizona Battery of Communication Disorders of Dementia – Verbal Learning subtest
AD	Alzheimer’s disease
ANOVA	Univariate analysis of variance
BG	Basal ganglia
BNT	Boston Naming Test – Second Edition
C-units	Communication units
CERAD	Consortium to establish a registry for Alzheimer’s disease.
CHAT	Codes for Human Analyses of Transcripts
CHILDES	Child Language Data Exchange System
CI	Confidence interval
CIU	Correct information units
COMT	Catechol-o-methyl transferase
Ctrl	Control(s)
DA	Discourse analyses
DAT	Dementia Alzheimer’s type
DCT	Discourse Comprehension Test
DiV	Denotes an individual dependent variable in a MANOVA analysis
DLB	Dementia with Lewy Bodies
DRS	Dementia Rating Scale
DV	Dependent variable
DcV	Denotes a dependent canonical variable in a MANOVA analysis
ES	Effect size

GPI	Globus pallidus interna
HC	Healthy control
HD	Huntington's disease
ICC	Intraclass correlation coefficient
IQ	Intelligence Quotient
IV	Denotes an independent variable in univariate F tests, t-tests, and discriminate function analysis
LBD	Lewy Body Dementia
LBSD	Lewy Body Spectrum Disorder. Superordinate
LED	Levodopa equivalent dose
MANOVA	Multivariate analysis of variance
MAO-B	Monoamine oxidase B
MDRS-2	Mattis Dementia Rating Scale – 2
MDS-UPDRS	Movement Disorders Society - Unified Parkinson Disease Rating Scale (revised version)
MLU	Mean length of utterance
MMSE	Mini Mental Status Examination
<i>n.s.</i>	non-significant statistical effect
PD	Parkinson disease
PD-MCI	Parkinson disease mild cognitive impairment
DFA	Discriminant function analysis
PDD	Parkinson disease dementia
PDND	Parkinson disease no dementia
RQ	Research question

SALT	Systematic Analysis of Language Transcription
SIT	Sentence Intelligibility Test
SLP	Speech-language pathologist
SNr	Substantia nigra pars reticulata
SPPT	Sentence Production Priming Test
T-units	Minimal terminal language unit (Hunt, 1965) a rule based system for segmenting spoken language
TAWF-verb	Test of Adolescent and Adult Word Finding
Trails	Trail Making Test
UPDRS	Unified Parkinson Disease Rating Scale – original version
UPDRS-III	UPDRS Part III motor assessment
WPM	Words per minute

Chapter 1: Thesis Overview

Parkinson disease is a progressive neurological disorder that affects both motor (e.g., movement, walking, speech) and non-motor (e.g., cognition, mood) functions with a significant impact on health-related quality of life. Dr. James Parkinson first described Parkinson disease (PD) in his seminal paper originally published in 1817: *An Essay On The Shaking Palsy* (Parkinson, 2002). Charcot further identified an akinetic form of the disease (i.e., non-tremulous) and in the late 19th century established the term Parkinson disease (i.e., “maladie de Parkinson”) in our medical lexicon (Jankovic, 2008). Dr. Parkinson described his initial six cases of PD as exhibiting resting tremor, impaired posture, and difficulty walking. Despite his detailed descriptions of the motor symptoms, Dr. Parkinson fell short of capturing fully the non-motor consequences of the disease particularly those related to cognition and to language. By documenting the observation in his patients that “the senses and intellects” were “uninjured” he would influence clinicians over the next 150 years toward the view that PD was exclusively a ‘motor’ disease (Parkinson, 2002).

PD is a complex and progressive degenerative neurological disease. While its hallmarks are dopaminergic dysregulation and basal ganglia dysfunction, PD also impacts other neurotransmitter systems affecting mood regulation and frontal lobe cognitive functions (Zgaljardic, Foldi, & Borod, 2004). Cognitive impairment in PD was recognized fully following the discovery of abnormal protein aggregates, Lewy bodies, in diseased brain cells in the early 1900’s and the further discovery of their relevance to cognitive impairment in the 1960’s (Holdorff, 2002; Okazaki, Lipkin, Aronson, 1961). Today, it is commonly accepted that cognitive changes begin early in PD (Elgh et al., 2009). More importantly, such changes can occur even in the absence of dementia (Holdorff, 2002). The downstream effect of PD-specific cognitive changes on language (e.g., understanding sentences, retrieving words) and communication (e.g., expressing ideas, maintaining relationships) is a rapidly developing area of research. However, work on these aspects has focused largely on more discrete language tasks (e.g., verbal fluency,

confrontation naming, and sentence processing) and motor speech tasks (e.g., voice intensity, articulation).

Disruptions in cognitive-linguistic functions can interact with motor speech impairments and can contribute to the communication difficulties that are of significant concern for individuals living with PD and for their care partners. In support of this assertion, Miller, Noble, Jones, and Burn (2006) reported that the communication challenges described by individuals with PD are multidimensional resulting not only from speech and voice changes but also from cognitive-linguistic symptoms such as: “formulating ideas”, “word retrieval”, “loss of train of thought”, “distractibility”, and “attention” (p. 237). Given the complexity of cognitive, language, motor speech, and gestural challenges in PD, the study of discourse production is ideal for revealing the potential interactions among these domains. Well-controlled spoken discourse tasks with systematically applied analyses are powerful tools in the discernment of cognitive-linguistic and motor speech impairments beyond the single word level (Cannizzaro & Coelho, 2013; Shadden, 1998a). However, to date, the impact of and integration of cognitive, language and motor speech challenges on spontaneous language tasks (i.e., discourse tasks) in PD remains relatively unexplored compared to more discrete aspects of cognition, language, and motor speech/voice.

1.1 Study Objective

The objective of the current study was to create a profile of spoken discourse impairment in PD that, with acceptable sensitivity and specificity, distinguished the spoken discourse of individuals with PD from the spoken discourse of healthy older controls.

1.2 Organization of Thesis Chapters

Chapter 2 begins with a brief overview of PD to increase the reader’s awareness of the magnitude and complexity of this disease including: the scope of its impact, diagnostic criteria, and typical management approaches. Following this overview, a cursory discussion of neurodegeneration in PD is provided to build a rationale for the importance of exploring cognitive-linguistic related processes as a downstream effect of the specific

neural changes in PD. Brief summaries of the cognitive and language impairments reported in the neuropsychological literature and of the motor speech consequences of PD are presented as a background for contextualizing the current body of PD discourse literature in relation to these more discrete impairments. The majority of *Chapter 2* (Section 2.3) is a comprehensive review of the current published literature exploring spoken monologic discourse performance in PD.

Given the focused objective of the current project on spoken discourse in PD, studies of conversational discourse and discourse comprehension in PD have been excluded purposefully. Also, because the current study focuses on spoken discourse production in idiopathic PD in the absence of dementia, unless a study or finding is of particular relevance, *Chapter 2* purposefully excludes those studies exploring cognitive, linguistic, motor speech, or discourse changes associated with: (a) hereditary forms of PD, (b) deep brain stimulation for the management of PD symptoms, (c) dementia with Lewy bodies (DLB), (d) Parkinson disease dementia (PDD), and (e) specific effects of ‘on’ vs. ‘off’ states of PD medication. In Section 2.4, a summary is presented of the major findings from the review of the literature on spoken discourse in PD. *Chapter 2* concludes with a rationale for the study (Section 2.5) and the hypothesis/research questions addressed in the current study (Section 2.6).

Chapter 3 addresses the study design and methodology details. *Chapter 4* follows with detailed description of the participants and a presentation of the results of the study. The interpretation of these results contextualized in the existing published literature, an analysis of the study limitations, and a brief discussion of the clinical relevance of the study is presented in *Chapter 5*.

Chapter 2: Background

2.1 Parkinson Disease: An Overview

In section 2.1 an overview of PD is presented to illustrate the complexity of this disease. A brief summary of PD including the magnitude of its impact, diagnostic criteria, and management challenges are presented. Following, a survey of the nature of and mechanisms of neurodegeneration in PD is provided to create a foundation for conceptualizing PD as a cognitive-linguistic disorder beyond the traditional boundaries of its motor impact.

Scope and management of PD. The scope of PD worldwide is substantial. The incidence rate (i.e., newly diagnosed cases each year) of PD in Canada is 20/100,000 people (Parkinson Society Canada, 2003). Moreover, the prevalence rate in North America is between 100 and 250/100,000 people resulting in almost 100,000 Canadians with PD (Parkinson Society Canada, 2003; Wolters & Bosboom, 2007b). While these statistics are based on North American data, they are comparable to the incidence and prevalence rates worldwide (Wolters & Bosboom, 2007b). According to the Parkinson Disease Foundation, 7 to 10 million people worldwide are living with PD (http://www.pdf.org/en/parkinson_statistics). The mean age of onset of PD is 62.4 years, whereas in Canada it is estimated that 85% of cases diagnosed are over the age of 65 (Parkinson Society Canada, 2003). The number of individuals with PD in both Canada and the United States is estimated to double by the year 2030 due to the aging population in these countries. In light of these numbers, individuals with PD represent a substantial and growing population requiring the expertise of communication disorders scientists and professionals.

As was the case in the eras of Dr. Parkinson and Dr. Charcot, even today PD is diagnosed primarily by clinical exam. The cardinal motor features by which a diagnosis of PD is made are bradykinesia plus the presence of at least *one* of the following: a) muscle rigidity, b) 4-6 Hz resting tremor, and c) postural instability not caused by other primary conditions (Hughes, Daniel, Kilford, & Lees, 1992; Jankovic, 2008). The diagnostic criteria for PD are presented below in Table 1.

Table 1

Diagnostic Features of Idiopathic Parkinson Disease (PD) (Gelb, Oliver, & Gilman, 1999; Hughes et al., 1992; Jankovic, 2008)

Core Features	Bradykinesia ¹ <i>And at least one of the following:</i> <ul style="list-style-type: none"> • Rigidity • 4 – 6 Hz rest tremor • Postural instability²
Supporting Features (3 or more required for diagnosis of PD)	<ul style="list-style-type: none"> • Unilateral onset • Tremor present only or more prominently with limbs at rest • Progressive disorder (clinical course ≥ 10 years) • Persistent asymmetry in motor symptoms • Excellent motor symptom improvement (70-100%) with levodopa • Severe chorea (dyskinesia) induced by levodopa treatment • Motor symptoms responsive to levodopa ≥ 5 years
Features Suggestive of other Disorders	<ul style="list-style-type: none"> • Other causes of parkinsonism³ such as focal brain lesion in a symptom-specific region, neuroleptic use within previous 6 months, and communicating hydrocephalus. • Repeated strokes resulting in parkinsonism • Repeated head injury result in parkinsonism • History of definite encephalitis • Hallucinations unrelated to medication during initial 3 years of diagnosis • Severe freezing of gait during initial 3 years of diagnosis • Severe dysautonomia unrelated to medications • Prominent postural instability early in the disease • Dementia preceding motor symptoms or occurring in the first year of diagnosis • Supranuclear gaze palsy • MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine) exposure

Note. ¹ Bradykinesia = Slowing of movement and progressive decrement of amplitude of movement during repetitive motor tasks such as foot tapping, finger tapping or hand opening and closing. ² Postural instability = not resulting from vestibular, neuropathy or other medical causes. ³ Parkinsonism = clinical term referring to a cluster of physical symptoms including slowing of movement, tremor, rigidity/stiffness, and balance problems resulting from a number of different brain pathologies.

The effective management of PD involves a multifaceted approach including pharmacotherapy, surgical interventions, and behavioral interventions (e.g., rehabilitation, counseling) (Martin & Wieler, 2003; Wolters & Bosboom, 2007b). The complexity of treating PD often is underestimated. The pharmacotherapy treatment of PD is complicated and becomes more challenging with disease progression (Martin & Wieler, 2003; van Laar, 2007). Multiple medications (i.e., carbidopa/levodopa, MAO-B inhibitors, dopamine agonists, COMT inhibitors) are now approved for the management of PD symptoms, although levodopa remains the most commonly used treatment (Schapira, 2007). Moreover, as the disease progresses and the occurrence of non-motor symptoms increase (e.g., gastrointestinal, dementia, mood) more intricate pharmaceutical management often is required. While some patients are candidates for surgical procedures (e.g., deep brain stimulation) the majority of patients are managed with pharmacotherapy. Deep brain stimulation is often limited to a select group of individuals with PD, requires ongoing monitoring and re-adjustment of stimulator settings by a programming expert and can have adverse effects on cognition (Esselink, de Bie, Schuurman, & Speelman, 2007; Halpern, Rick, Danish, Grossman, & Baltuch, 2009). Behavioural interventions (e.g., speech, voice, and language therapy; physiotherapy; occupational therapy; therapeutic recreation) are based on the needs of patients and caregivers and change over the continuum of the disease necessitating ongoing re-evaluation of goals and therapeutic interventions.

While the diagnostic and management criteria generally are well established for PD, the etiology and pathogenesis of the disease remain strongly debated (Bartels & Leenders, 2009; Olanow & Tatton, 1999; Wolters & Bosboom, 2007b). The commonly accepted consensus is that PD is caused by a combination of genetic and environmental influences often referred to as the “double hit hypothesis” (Bartels & Leenders, 2009; Olanow & Tatton, 1999). Almost 85% of PD cases are ‘idiopathic’ meaning that no exact cause for the development of the disease is known.

Neurodegeneration in PD. Important to the understanding of cognitive-linguistic impairments in PD is the recognition that, although PD primarily affects the basal ganglia (BG), the disease is not restricted to subcortical structures in its impact

(Aarsland, Perry, Brown, Larsen, & Ballard, 2005; Bartels & Leenders, 2009; Braak et al., 2003; Braak, Ghebremedhin, Rub, Bratzke, & Del Tredici, 2004; Braak & Del Tredici, 2008; Del Tredici, Rub, De Vos, Bohl, & Braak, 2002; Wolters & Bosboom, 2007b). Although the study of brain dysfunction in PD has advanced immensely over the last 40 years, much controversy exists regarding the exact mechanisms of neuronal injury and their correlation to motor and non-motor symptoms associated with the disease (Bartels & Leenders, 2009; Olanow & Tatton, 1999; Wolters & Bosboom, 2007b). However, there is general consensus regarding two primary mechanisms of injury leading to PD: (a) disruption of critical neurotransmitters (i.e., how brain cells facilitate information flow) including but not exclusively dopamine and (b) neuronal cell dysfunction caused in part, although not exclusively, by the aggregation of Lewy bodies, Lewy neurites, and alpha synuclein (Bartels & Leenders, 2009; Wolters & Bosboom, 2007b). While it is beyond the scope of this chapter to discuss the neurodegeneration aspects in detail, a brief overview follows of how these mechanisms specifically alter cortical and subcortical functioning and the subsequent impact on cognitive-linguistic functions.

The role(s) of the BG in cognition and language have long been under-appreciated because of scientific and clinical biases toward a cortical-centric view of language and a motor-centric view of the BG. However, recent advances in neuroimaging techniques (e.g., tractography imaging, connective analysis methods), neurosurgical recordings from deep brain structures and retrograde transneuronal tracers have expanded our knowledge of the role of the BG in cognition and language (Barbas, Garcia-Cabezas, & Zikopoulos, 2013; Carbon & Marie, 2003; Middleton & Strick, 2000). The BG is a series of bilateral, interconnected, and heterogeneous subcortical grey matter structures that are intricately connected to both the cortex and the cerebellum (Alexander & Crutcher, 1990; Carbon & Marie, 2003; Groenewegen, 2003; Herrero, Barcia, & Navarro, 2002; Middleton & Strick, 2000). Within the scope of this chapter, the discussion is limited to the relationship of the BG with the left prefrontal cortex because of the critical roles played by the left prefrontal cortex in the discrete cognitive processes that sub-serve language. However, cognitive-linguistic processes are also mediated by the connections between the BG and right prefrontal cortex although to a lesser degree. The BG have a

preferential relationship with the prefrontal and motor cortices by which they modulate activity in the cortex via the basal ganglia-thalamic-cortical pathways (Alexander, DeLong, & Strick, 1986; Groenewegen, 2003). This preferential relationship includes connections with cortical centers for language and speech production such as Broca's area and the dorsolateral prefrontal cortex, regions critical for cognitive processes that buttress language functions (Barbas et al., 2013; Carbon & Marie, 2003; Chenery, Angwin, & Copland, 2008; Crosson, 1985; Crosson, ; Duffau et al., 2003; Duffau et al., 2005; Duffau, 2008; Ellmore, Beauchamp, O'Neill, Dreyer, & Tandon, 2009; Groenewegen, 2003; Herrero et al., 2002; Middleton & Strick, 2000; Ullman, 2006). For an excellent review of this topic regarding the relationship among the basal ganglia and traditionally viewed cortical language centers see Barbas et al. (2013).

Although the number of circuits and their paths vary, research from both human and animal models suggest that the reciprocal communication between the BG and the cortex is accomplished via predominately parallel, behaviorally specialized, and somatotopically specific circuits (Alexander, Crutcher, & DeLong, 1990; Groenewegen, 2003; Herrero et al., 2002; Middleton & Strick, 2000). While these circuits are largely segregated from their point of origin in the cortex and remain separate as they course through the BG, research suggests that they may interact at three levels in the BG: the globus pallidus and substantia nigra pars reticulata (SNr) complex, the thalamus, and within the collateralized axons of the various BG nuclei (Herrero et al., 2002). This BG-specific organization structure may partially explain the presence of overlapping cognitive, language, and motor impairments in PD (Middleton & Strick, 2000).

Adding to the complexity of these interconnecting systems is the flow of information within the BG circuitry. Information flow is accomplished via a series of direct and indirect pathways among the cortex, the thalamus, and the BG (Alexander et al., 1990; Alexander & Crutcher, 1990; Barbas et al., 2013; Chenery et al., 2008; Groenewegen, 2003; Herrero et al., 2002). Briefly, the normal balance of activity in the direct pathway (i.e., inhibitory input to globus pallidus interna/SNr complex (GPi/SNr) resulting in cortical excitation) and in the indirect pathway (i.e., excitatory input to GPi/SNr resulting in cortical inhibition) facilitates normal motor and cognitive-linguistic functions as

modulated by the neurons in the cortex and the BG (Alexander & Crutcher, 1990; Barbas et al., 2013; Chenery et al., 2008; Groenewegen, 2003; Herrero et al., 2002).

Consequently, PD-specific disruption in the balance of the excitatory and inhibitory pathways, largely via disruption of the neurotransmitter dopamine, alters both motor and cognitive functions served by the corticobasal circuitry (Alexander & Crutcher, 1990; Barbas et al., 2013; Chenery et al., 2008; Groenewegen, 2003; Herrero et al., 2002).

While, the specific effects of this imbalance in PD continues to be investigated, currently it is hypothesized that the disease results in a marked increase in GPi/SNr activity that subsequently increases the inhibition of the thalamic output and thus reduces net cortical excitation resulting in motor and cognitive impairments (Herrero et al., 2002).

Importantly, the loss of dopamine and dopaminergic receptors in PD is not homogenous across the BG (Hornykiewicz, 2001). This heterogeneous pattern of cell loss partially accounts for the variability in motor and cognitive symptoms observed across individuals with PD (Hornykiewicz, 2001). In summary, PD-specific disruption of the normal buttressing provided to the cortex by the BG structures and the associated dopaminergic pathways provides a rationale for expanding our exploration of communication disorders among persons with PD beyond the impact of motor speech dysfunction to include the role of cognitive-linguistic dysfunction.

2.2 Cognition, Language, and Motor Speech Impairments

Section 2.2 summarizes briefly the existing literature relative to cognitive impairments in PD including a discussion of the scope of these challenges and a survey of cognitive impairments observed in the absence of dementia. To build a foundation for the discussion of the discourse literature in PD, an overview of the affected language, motor speech, and voice processes is then presented.

Cognition impairments in PD. Although not ubiquitous, for many individuals with PD cognitive changes are present from the earliest stages of symptomatic disease (Aarsland, Beyer, & Kurz, 2008; Aarsland et al., 2010; Benito-Leon et al., 2011; Elgh et al., 2009). In a longitudinal population based study conducted in Sweden, 30% of newly diagnosed persons with PD presented with cognitive impairment (Elgh et al., 2009).

Point prevalence data indicate that the cumulative prevalence of Parkinson disease

dementia (PDD) is as high as 78% at 8 years post-onset of PD (Aarsland, Andersen, Larsen, & Lolk, 2003). The relative risk of developing dementia is 4 to 6 times higher for individuals with PD compared to community dwelling healthy adults (Emre et al., 2007). PDD is associated with older age at onset of disease, increased falls, increased gait and balance instability, and rapid eye movement (REM) behaviour sleep disorder (Emre et al., 2007).

Multiple studies have profiled the cognitive deficits in PD using neuropsychological measures with robust and consistent findings (Aarsland et al., 2010; Benito-Leon et al., 2011; Elgh et al., 2009; Muslimovic, Post, Speelman, & Schmand, 2005; Pagonabarraga & Kulisevsky, 2012; Rodriguez-Ferreiro, Cuetos, Herrera, Menendez, & Ribacoba, 2010; Wolters & Bosboom, 2007a). The consistently reported domains of impairment in the early stages of PD in the absence of dementia include:

- Impaired verbal fluency for semantic categories, phonemes and action words (Benito-Leon et al., 2011; Elgh et al., 2009; Green et al., 2002; Henry & Crawford, 2004; Muslimovic et al., 2005; Pagonabarraga & Kulisevsky, 2012; Rodriguez-Ferreiro et al., 2010);
- Impaired attention (e.g., activation, sustained, set-shifting) (Elgh et al., 2009; Green et al., 2002; Muslimovic et al., 2005; Pagonabarraga & Kulisevsky, 2012);
- Impaired working memory (Muslimovic et al., 2005);
- Impaired psychomotor speed (Elgh et al., 2009; Muslimovic et al., 2005);
- Impaired verbal and visual memory (Benito-Leon et al., 2011; Bohlhalter, Abela, Weniger, & Weder, 2009; Elgh et al., 2009; Ivory, Knight, Longmore, & Caradoc-Davies, 1999; Muslimovic et al., 2005; Pagonabarraga & Kulisevsky, 2012);
- Impaired visual construction/visual perceptual skills (Pagonabarraga & Kulisevsky, 2012).

Cognitive impairment occurring early in PD is not a precursor per se to PDD (Svenningsson, Westman, Ballard, & Aarsland, 2012; Wolters & Bosboom, 2007a). In fact, the rate of cognitive decline is quite variable among individuals with PD. Typically, the rate of decline in cognition is slower during the first 5 to 7 years of PD and then

progresses more rapidly afterwards (Gago et al., 2009; Svenningsson et al., 2012; Wolters & Bosboom, 2007a). Additionally, the pace of cognitive decline generally mirrors the pace of and severity of motor decline. Interestingly, rapid declines in speech intelligibility, increases in axial motor symptoms, and declines in gait are associated with faster rates of cognitive decline and development of PDD (Burn et al., 2003; Elgh et al., 2009; Gago et al., 2009; Green et al., 2002; Verbaan et al., 2007; Locascio, Corkin, & Growdon, 2003; Sollinger, Goldstein, Lah, Levey, & Factor, 2010). The degree of decline across cognitive processes also is not homogenous. The mechanisms of this variability in cognitive impairment in PD are not clear and continue to be investigated. However, it is suggested by some researchers that those individuals with cognitive impairments involving exclusively the fronto-striatal circuitry (e.g., pathways that support executive function and working memory) remain more stable over the course of the disease with less risk of developing dementia (Green et al., 2002; Henry & Crawford, 2004; Pagonabarraga et al., 2008; Pagonabarraga & Kulisevsky, 2012; Svenningsson et al., 2012). Conversely, those individuals with PD who also demonstrate extra-prefrontal cortical deficits such as constructional deficits (e.g., parietal lobe), semantic, and language impairments (e.g., temporal lobe) may reflect a more wide-spread cortical involvement with a greater likelihood of developing dementia (Aarsland et al., 2005; Braak, Rub, Jansen Steur, Del Tredici, & de Vos, 2005; Green et al., 2002; Pagonabarraga & Kulisevsky, 2012; Svenningsson et al., 2012). Collectively, findings from these studies highlight the heterogeneous nature of cognitive decline in PD and its relationship to the underlying complexity of neurodegeneration in PD. Moreover, these findings underscore the potential value of using discourse tasks, which require the coordinated interaction of cognitive, language, and motor speech processes (Ulatowska, Chapman, Bloom, & Obler, 1994), for revealing subtle patterns of and changes in cognition in PD.

Language impairments in PD. While BG structures and pathways have been implicated in language-specific functions, the role(s) of the BG in language comprehension and production remain unresolved. The consensus among researchers is that language deficits in disorders associated with the disruption of BG circuits (e.g. PD, Huntington's disease, subcortical stroke) likely result from an impairment in cognitive

processes that are essential for, indeed integrated with, successful language comprehension and production (Barbas et al., 2013; Chenery et al., 2008; Crosson, 1985; Crosson, ; Nadeau, 2008). In fact, many of the cognitive processes associated with language functions (e.g., working memory, attention allocation, selection processes, inhibition processes, error detection, error resolution, rule-based learning, and sequencing) have been attributed to the dopamine-mediated BG circuits and the prefrontal cortex. Whether or not the BG has specific roles in language processing and production independent of cognitive processes remains a key question.

Language impairments in persons with PD, while not ubiquitous, are reported frequently in the more recent literature. The impairments include disrupted word retrieval in verbal fluency (i.e., phonemic, semantic, action, and alternating fluency tasks) (Henry & Crawford, 2004) and impaired confrontation naming (i.e., naming of pictured objects and verbs) (Cotelli et al., 2007; Herrera, Rodriguez-Ferreiro, & Cueto, 2012; Rodriguez-Ferreiro, Menendez, Ribacoba, & Cueto, 2009). While there is general agreement regarding the existence of verbal fluency deficits in PDD, the presence of verbal fluency deficits in PD without dementia is not as clear. Further, performance on verbal fluency in PD varies substantially across different tasks (e.g., phonemic fluency vs. semantic fluency). However, a recent meta-analysis suggests that individuals with PD are more impaired on semantic and phonemic verbal fluency tasks vs. healthy controls (Henry & Crawford, 2004). Adding to the complexity of word retrieval problems in PD, there is emerging but consistent evidence that individuals with PD exhibit a specific deficit in word retrieval for verbs (Cotelli et al., 2007; Crescentini, Mondolo, Biasutti, & Shallice, 2008; Fernandino et al., 2012; Herrera & Cueto, 2012; Herrera et al., 2012). Crescentini, Mondolo, Biasutti, and Shallice (2008) recently hypothesized that the verb deficit in persons with PD is not a specific language impairment of the semantic system per se but rather is a deficit of the supervisory or attention mechanisms for lexical selection. However, others interpret the presence of verb-specific deficits in PD as a deficit in the access to and retrieval of semantic and lexical information relating to action language including the motion-related features of those semantic concepts. Such an interpretation suggests that there is a relationship between movement perception in PD and language processes; an interpretation that is supported in both behavioural and

imaging data (Boulenger et al., 2008; Herrera et al., 2012; Peran et al., 2009; Peran et al., 2010; Pulvermuller, 2005; Pulvermuller, Shtyrov, & Ilmoniemi, 2005). Regardless of the mechanism(s) of impairment in verb production in PD, deficits in the retrieval of verbs are likely to interfere significantly with communication among persons with PD and their caregivers. Furthermore, such verb processing and retrieval deficits may interact with other linguistic processes such as sentence comprehension and production, both of which are highly dependent on the ability to access the semantic and lexical information from verbs (Hare, Jones, Thomson, Kelly, & McRae, 2009).

One of the more extensively explored areas of language in persons with PD is their impaired processing of syntax and grammar. Multiple investigators have demonstrated that individuals with PD are less accurate and/or less efficient vs. controls on tasks requiring extracting information and decoding syntax from complex sentence structures (e.g., center embedded structures, non-canonical word order structures, and semantically non-constrained sentences) that either are heard or are read (Grossman et al., 1991; Grossman, Crino, Reivich, Stern, & Hurtig, 1992b; Grossman, Carvell, Stern, Gollomp, & Hurtig, 1992a; Grossman, Carvell, & Peltzer, 1993; Grossman et al., 2012; Hochstadt, Nakano, Lieberman, & Friedman, 2006; Hochstadt, 2009; Natsopoulos et al., 1993; Skeel et al., 2001). Moreover, individuals with PD show consistent impairments in comprehension of sentences resulting from impaired processing of semantic ambiguities (For review see Chenery et al., 2008). Collectively, the studies of syntax processing among individuals with PD suggest that they experience challenges disambiguating syntax and semantics to extract meaning from sentences. These deficits are magnified by PD-specific impairments in frontostriatal-mediated executive functions. Researchers demonstrated further that individuals with PD exhibit more challenges processing sentences in the presence of increasing cognitive loads originating either within the inherent nature of the linguistic task or within the demands of the environment in which sentences are processed (Grossman et al., 2000; Grossman et al., 2003). These findings suggest that studying and understanding cognitive-linguistic abilities of individuals with PD in dynamic and contextual language tasks are of paramount importance.

Compared to the comprehension domain, there is scant published literature on syntax production in persons with PD. In a recent publication by Troche and Altmann (2012) English-speaking participants with PD produced sentences using picture stimuli that were designed to elicit complex syntax structures varying across two levels of complexity based on the number of verb clauses. In the less complex syntax condition, participants generated one-event sentences (e.g., “The passenger from the suburbs waited for the train.”) whereas in the more complex condition they generated two-event sentences (e.g. “The bully frightened the girl who jumped over the fence.”) (Troche & Altmann, 2012, p. 234). The authors reported that participants with PD produced sentences that were judged as less grammatically correct and less complete vs. those of controls. Furthermore, participants with PD experienced higher rates of verbal disruption (i.e., a signal of potential challenges in cognitive, language, and/or motor planning) than controls, a finding that was exaggerated in the more complex condition. Colman et al. (2009) studied verb use in spontaneously generated sentences in a group of Dutch-speaking participants with PD and healthy controls (Colman et al., 2009). Their PD participants produced significantly more errors on present tense vs. past tense verbs and on intransitive vs. transitive verbs. While the number of studies is limited, collectively these studies conducted across two different languages suggest that grammar production is problematic in PD, which if manifested in the context of spoken discourse would substantially affect communication.

The term figurative language refers to words or groups of words that are interpreted using implicit or non-defined meanings. Specifically, figurative language includes, but is not limited to, metaphors, hyperbole, counterfactual statements, proverbial statements, irony, and similes. In healthy individuals, figurative language is complex and depends on a myriad of cognitive-linguistic processes including, but not limited to, accessing the literal meaning of words, processing the perspective of the speaker/writer (i.e., theory of mind), determining the context around the utterance, processing the non-verbal information from prosody/punctuation, and discovering the intent or goal of the message. While still a relatively unexplored area of language in persons with PD, researchers revealed a variety of deficits involving figurative language comprehension and production. Researchers showed consistently that individuals with PD experience challenges interpreting the

intention or goal of a passage/statement (i.e., inferencing) (Berg, Bjornram, Hartelius, Laakso, & Johnels, 2003; Bhat, Iyengar, & Chengappa, 2001; Holtgraves & McNamara, 2010; Lewis, Lapointe, Murdoch, & Chenery, 1998; Monetta, Grindrod, & Pell, 2009; Murray & Stout, 1999). The work of Monetta, Grinrod, and Pell (2009) suggests that individuals with PD exhibit challenges specifically interpreting irony statements based on difficulties determining the mental perspective of the characters in a story suggesting potential theory of mind deficits. Studies also demonstrated PD-specific impairments in producing and comprehending counterfactual statements (McNamara, Durso, Brown, & Lynch, 2003; Monetta et al., 2009). Counterfactual statements as defined by McNamara et al. (2003) are “mental representations of alternatives of past events” or “imagined alternatives to something that has actually occurred” (McNamara et al., 2003, p. 1065). Increasing the complexity of figurative language deficits in PD, investigators reported consistently that persons with PD are more impaired vs. healthy controls in extracting accurately the meaning from metaphorical language, such as “You can’t have your cake and eat it too” (Berg et al., 2003; Lewis et al., 1998; McKinlay, Dalrymple-Alford, Grace, & Roger, 2009; Monetta & Pell, 2007; Monetta et al., 2009).

Collectively, the body of literature relative to cognitive-linguistic impairments in PD provides unambiguous evidence that supports the hypothesis that impairments in cognition and language can manifest in discourse tasks. More importantly, the evidence also reveals that spontaneous language tasks may inform our understanding of the challenges in communication reported by individuals with PD and their caregivers that can facilitate the development of language and communication intervention strategies.

Motor speech and prosody impairments in PD. The impact of cognitive and language impairments on discourse in PD is complicated further by the prominent presence of motor speech problems typically presenting as hypokinetic dysarthria. Upwards of 90% of individuals with PD will develop articulation and/or voice changes such as reduced precision of articulation, changes in speaking rate (i.e., either increasing or decreasing rate), reduced loudness, and breathy/hoarse voice quality that can affect communication at the discourse level (Darley, Aronson and Brown, 1975; Ramig, Fox, & Sapir, 2004). PD-specific neurological changes result in challenges ‘scaling’ movements

and determining/judging the appropriate degree of motor response to a specific task requirement (Pinto et al., 2004; Ramig et al., 2004). However, more importantly other communication-related changes in PD such as reduced pitch variability, reduced stress patterns, and reduced emotional expression via vocal, gestural, and facial channels can interact substantially with the ability to convey meaning through language in discourse (Pell, Cheang, & Leonard, 2006; Pitcairn, Clemie, Gray, & Pentland, 1990; Ramig et al., 2004). The disruptions in prosody are not limited to the production domain among persons with PD but also are present in the comprehension domain, which can reduce the effectiveness of individuals' roles of speaker and listener (Monetta, Cheang, & Pell, 2008; Pell et al., 2006). Importantly, motor speech and prosodic impairments in PD contribute to negative listener perceptions including the view that individuals with PD are less interested in communication, less involved in conversation, less happy, less friendly, and more anxious (Jaywant & Pell, 2010; Pitcairn et al., 1990).

2.3 Studies of Spoken Monologic Discourse in PD

The cognitive, language, and motor speech impairments reported in the PD literature underscore the importance of studying the impact of these deficits on discourse performance. Within the field of discourse research, it is a widely held assumption that spoken discourse production reflects the integration of cognition, language, and motor speech abilities. Moreover, different discourse tasks/genres present differing levels of cognitive and/or language demands. Spoken discourse production tasks offer a unique opportunity to advance our understanding of how discrete impairments of cognition (e.g., attention switching), language (e.g., production of complex syntax structures), and motor speech (e.g., speech intelligibility) are manifest in spontaneous spoken language tasks. Given the complexity of cognitive, language, and motor speech challenges in persons with PD, the study of monologic discourse production presents an ideal format for revealing the potential interactions among and the functional impacts of these changes within a spontaneous but relatively structured language production task.

The remainder of this chapter focuses on studies that use discourse analyses with tasks of spoken monologic discourse production in individuals with idiopathic PD without

dementia. Sections of 2.3 have been reproduced in part and or edited with permission of the copyright holders from Roberts and Orange (2013).

Monologic discourse defined. Discourse as defined by Ulatowska and Olness (2004), and the definition used in this dissertation, is language (e.g., spoken, written, gestured, or signed) that is “beyond the boundaries of isolated sentences” (Ulatowska & Olness, 2004, p. 300). Monologic spoken discourse tasks refer to those typically high-structured discourse tasks (e.g., picture description, procedural narratives) designed to elicit samples of spontaneously produced language. While monologic discourse tasks reflect the demands of spontaneous language during every day conversational tasks (Doyle, Goda, & Spencer, 1995), they are not, by their nature, interactional. Unlike conversational discourse, monologic discourse tasks do not per se engage an individual in the active, dynamic, and dyadic exchange of novel information where partners have potentially equal roles in the exchange. Furthermore, some monologic discourse tasks, while representative of everyday language, do not serve the purpose of building or maintaining relationships, as conversational discourse does. However, the structured nature of monologic discourse tasks that enable manipulation and control of cognitive-language variables make them uniquely valuable for the study of the interactions among cognition, language, and motor speech components of communication; giving these tasks special relevance for the assessment and interpretation of discourse performances among persons with PD.

Commonly used monologic discourse tasks include picture description (i.e., single pictures and picture sequences), narrative generation, story retelling (i.e., retelling a story from memory), procedural descriptions (i.e., describing the steps for completing a specified task), and personal narratives including topic guided/directed interviews. While personal narrative tasks (i.e., the use of structured questions to elicit discourse samples around personal information such as health, occupation, family, etc.) can be more ‘conversation like’ and ‘dyadic’ in nature, for this review they are included as monologic tasks. In the studies reviewed, personal narrative tasks were used to elicit extended spoken discourse samples and did not per se involve the exchange of novel information

between equal partners (i.e., the participant and the examiner). As such, the task was used in a more monologic discourse sense vs. conversation sense.

Each of these monologic tasks presents a unique set of cognitive and linguistic demands (e.g., memory, personal relevance, linguistic specificity, syntactic complexity, etc.) that can affect discourse performance (Shadden, Burnette, Eikenberry, & DiBrezzo, 1991). The use of monologic discourse tasks is well established in the literature in cognitive-communication disorders resulting from neurodegenerative processes (Ash et al., 2006; Orange & Kertesz, 2000; Roberts-South, Findlater, Strong, & Orange, 2012; Zraick, Carr, Gregg, & Smith-Olinde..., 2011). For example, monologic discourse tasks can reveal subtle changes in cognitive-linguistic functions in mild cognitive impairment (Fleming & Harris, 2008). Multiple monologic discourse tasks also can be used to reveal participants' abilities and impairments in cognition, language, and motor speech production as a function of the intrinsic demands of the task (e.g., high vs. low lexical specificity, high vs. low recall demands, high vs. low sequencing demands) (Shadden, 1998a).

Table 2 contains a compendious summary of the published literature reporting performances on spoken monologic discourse tasks among persons with PD including: (a) study authors and year of publication, (b) the discourse genre, (c) types of discourse analyses conducted, and (d) relevant demographic data for the participant groups. Table 2 reveals the large diversity in methods, tasks, analyses, and participants used to explore discourse production in PD to date. While informative in its breadth, the diversity of methods presents challenges in comparing results across studies.

Table 2

Summary of Published Discourse Production Studies in Parkinson Disease (Edited from (Roberts & Orange, 2013))

Study	Discourse Genre/Analyses	Participants
Ash, et al. (2011)	<ul style="list-style-type: none"> • Narrative discourse (story generation) using the wordless picture book <i>Frog Where Are you?</i> by Mercer Mayer • Analyses: Microlinguistic, macrostructural 	<ul style="list-style-type: none"> • N = 32 LBSD (n = 14 PDD/DLB; n = 18 PDND) & N = 16 HC • Mean age = 71.9(8.5) LBSD; 68.6(6.8) HC • Mean MMSE = 20.9(4.7) PDD/DLB; 27.9(1.5) PDND; 29.1(1.2) HC • Mean disease duration = 6.6(2.3) PDD/DLB; 6.1(3.2) PDND • Mean UPDRS-III (/108) = 24.1(13.3) PDD/DLB; 20.9(9.1) PDND • Mean Hoehn & Yahr (/5) = 2.7(0.6) PDD/DLB; 2.2(0.7) PDND • LED = 287 (387) PD/DLB; 534 (376) PDND • Speech intelligibility data not provided • No depression screen • No measure of IQ
Ash, et al. (2012)	<p>Narrative discourse (story generation) using the wordless picture book <i>Frog Where Are you?</i> by Mercer Mayer</p> <p>Analyses: Microlinguistic</p>	<ul style="list-style-type: none"> • N = 35 LBSD (n = 14 PDD/DLB; n = 21 PDND) & N = 16 HC • Mean age = 72(8.4) LBSD; 68.6(6.8) HC • Descriptives for PDD/DLB & HC groups identical for Ash et al. (2011). See above. • Mean MMSE 27.4(2.0) PDND • Mean disease duration = 6.0(3.1) PDND • Mean UPDRS-III (/108) = 21.7(9.0) PDND • Mean Hoehn & Yahr (/5) = 2.2(0.6) PDND • LED = 592 (373) PDND • Speech intelligibility data not provided • No depression screen • No measure of IQ
Bayles (1990)	<ul style="list-style-type: none"> • Immediate story-retelling using the Shopping story from the Arizona Battery of Communication Disorders in Dementia • Analyses: Macrostructural 	<ul style="list-style-type: none"> • N = 12 PDND; 32 HC • Mean age = 68(9.7) PDND; 70.25(8.66) HC • Cognitive assessment = Mental Status Questionnaire (Goldfarb & Antin). PDND and HC all scored between 8 - 10/10 reported as 'normal' cognition by author • Descriptives for motor severity, disease duration, disease staging, speech intelligibility, LED not provided • No depression screen • No measure of IQ

Study	Discourse Genre/Analyses	Participants
Godbout & Doyon (2000) ¹	<ul style="list-style-type: none"> • Procedural discourse • Analyses: Macrostructural 	<ul style="list-style-type: none"> • N = 16 PDND; 16 HC • Mean age = 59(5.6) PDND; 54.3(13.7) HC • PD subtype = 6 akinetic rigid; 2 tremor; 8 mixed • Cognitive assessment = C.E.R.A.D protocol of neuropsychological measures (Morris et al., 1989). Authors report as normal for PDND group. • Mean MMSE = 28.34(1.04) PDND; No score provided for HC • Range disease duration = 1 - 21 years • Range Hoehn & Yahr (/5) = 1 - 3 • Descriptives for motor severity, speech intelligibility, LED not provided • Depression screened by clinical record only • No measure of IQ
Huber & Darling (2011) ¹	<ul style="list-style-type: none"> • Expository discourse speaker's topic of choice • Analyses: Microlinguistic 	<ul style="list-style-type: none"> • N = 14 PD; 14 HC • Mean age = 69 PD women; 76 PD men; 71 HC women; 72 PD men • MMSE given but criteria described as 'pass' with no cut-off score provided • Range disease duration = 1 - 10 years • Range Hoehn & Yahr = 2 - 3 • Perceptual ratings of speech severity made by trained (SLP) listeners provided in article • Descriptives for motor severity. LED and cognitive measures not provided • No depression screen • No measure of IQ
Illes, Metter, Hanson, & Iritani (1988)	<ul style="list-style-type: none"> • Personal narrative discourse using topic directed interview format • Analyses: Microlinguistic 	<ul style="list-style-type: none"> • N = 10 PD (all men); 10 HC • Disease severity reported by <u>Webster 30-point Scale of Parkinsonian Disability</u> n = 5 mild severity; n = 5 moderate severity • Mean disease = 4.6 years mild group; 10 years moderate group • Cognitive function = self-report of change • Descriptives for age, Hoehn & Yahr, speech intelligibility, LED not provided • No depression screen • No measure of IQ

Study	Discourse Genre/Analyses	Participants
Illes (1989)	<ul style="list-style-type: none"> • Personal narrative discourse using topic directed interview format • Analyses: Microlinguistic 	<ul style="list-style-type: none"> • N = 10 PD; 10 HD; 10 DAT; 10 HC (all men) • Participants stratified into early or middle stages for respective diseases n= 5 mild; 5 moderate per subgroup • PD stratified based on Webster scale used in Illes et al. (1988) • Mean age = 56.4 PD early; 66.0 PD middle • Cognitive function = self/care provider report of no issues • Descriptives for HC ages, disease duration, motor severity, speech intelligibility, LED not provided • No depression screen • No measure of IQ
Jaywant & Pell (2010) ¹	<ul style="list-style-type: none"> • Picture description stimuli from <u>Discourse Comprehension Test</u> (Brookshire & Nicholas, 1997) • Analyses: Microlinguistic; Macrostructural (listener perceptions of content) 	<ul style="list-style-type: none"> • N = 18 PD; 17 HC • Mean age = 60.28(7.26) PD; 61.29(9.29) HC • Mean disease duration = 3.57(1.84) • Mean UPDRS-III(/108)= 13.67(7.3) • Mean Hoehn & Yahr = 1.75(0.46) • Mean Mattis Dementia Rating Scale (/144) = 141(2.12) PD; 142.41(1.87) HC • Descriptives for speech intelligibility and LED not provided • No depression screen • No measure of IQ
Mcnamara, Obler, Au, Durso and Albert (1992)	<ul style="list-style-type: none"> • Picture description Cookie Theft stimuli from <u>Boston Diagnostic Aphasia Exam</u> (Goodglass & Kaplan, 1983) • Analyses: Microlinguistic; self-monitoring and repair strategies 	<ul style="list-style-type: none"> • N = 178; n = 22 PD; 15 DAT; 141 HC intentionally sampled from 4 age groups (30's to 70's). The 37 participants in the 60's group corresponded in age to the PD group • Mean age = 61.3(2.9) PD; 64.4(2.9) 60's HC • Range Hoehn & Yahr = 2 - 3 • Cognitive assessment = reported as non-demented by authors; metric not provided • Descriptives for motor severity, disease duration, speech intelligibility, LED, and cognitive measures not provided • No depression screen • No measure of IQ

Study	Discourse Genre/Analyses	Participants
Murray (2000)	<ul style="list-style-type: none"> Picture description (2 discourse samples) Cookie Theft stimuli from the <u>Boston Diagnostic Aphasia Exam</u> (Goodglass & Kaplan, 1983) & Grocery Scene (Helm-Estabrooks, 1992) Analyses: Microlinguistic; Macrostructural 	<ul style="list-style-type: none"> N = 10 PD; 10 HD; 18 HC (9 matched to the HD group & 9 matched to the PD group) Mean age = 70.1(8.8); 67.89(6.13) PD-HC Mean disease duration = 6.4(2.91) Mean Hoehn & Yahr = 3 Mean Dementia Rating Scale (/144) = 132.7(6.82) PD; 138.67(3.35) PD-HC Speech intelligibility scores using standardized measures are reported in the article Descriptives of motor severity and LED not provided Formal screening measure for depression completed IQ estimate data provided
Murray & Lenz (2001)	<ul style="list-style-type: none"> Personal narratives using topic directed interview format Analyses: Microlinguistic 	<ul style="list-style-type: none"> N = 10 PD; 9 HD; 17 HC (9 matched to the HD group & 8 matched to the PD group) Mean age = 70.1(8.8); 68.9(5.7) PD-HC Mean disease duration = 6.4(2.9) Mean Dementia Rating Scale (/144) = 132.8(6.8) PD; 139.0(3.4) PD-HC Speech intelligibility scores using standardized measures are reported in the article Descriptives of motor severity, Hoehn & Yahr and LED not provided Formal screening measure for depression completed IQ estimate date provided

Note: ¹ = Studies in which discourse data were reported but for which the theoretical premise of the study was not the analysis of discourse. Age and duration data are presented in years. PD = Parkinson disease assumed or reported as having no dementia. PDD = Parkinson disease dementia. DLB= dementia with Lewy bodies. LBSD = Lewy body spectrum disorder comprised of individuals with PD, PDD, LBD. HC = Health controls assumed to be age-matched to disease group unless otherwise indicated. DAT = dementia Alzheimer's type. HD = Huntington's disease. UPDRS-III = Unified Parkinson Disease Rating Scale motor items. MMSE = Mini Mental Status Examination. Hoehn & Yahr = measure of symptom staging in Parkinson disease. LED = levodopa equivalent dose (mg). IQ = intelligent quotient. PDND = Parkinson disease no dementia. SLP = Speech-language pathologist.

To facilitate the review of the diverse body of literature on spoken monologic discourse performance among persons with PD, the studies and their results are grouped by discourse analyses instead of by discourse genre (i.e., task). To synthesize and to organize the relevant findings from these studies, the taxonomy for discourse analyses used by Cherney (1998) and Coelho (2007) was adopted as a framework for discussion of the existing literature relative to spoken monologic discourse tasks in PD: (a) microlinguistic (i.e., within-sentence level analyses), (b) microstructural (i.e., across-sentence level analyses), (c) macrostructural (i.e., analyses of thematic unity within the narrative, and (d) superstructural (i.e., story grammar level analyses). Within the published literature in PD, most of the analyses methods conducted to date can be categorized as either microlinguistic or macrostructural thus the review is limited to these two categories.

Microlinguistic analyses. The following is a discussion of the nascent studies and their findings from the spoken monologic discourse literature in PD. The sections are presented relative to within-sentence level analyses (i.e., productivity, syntax accuracy and complexity, lexical diversity, and verbal fragmentation analyses).

Productivity. Productivity measures in discourse reflect the amount of and efficiency of language produced in a sample often using the measurable units of words/syllables produced. Productivity measures are influenced highly by cognitive impairments, language disorders, and motor speech problems, among other factors (e.g., hearing). Jaywant and Pell (2010), using a picture description task, reported that individuals with PD produced discourse samples that were significantly shorter in duration vs. healthy controls (HC) (PD $M = 28.03$ sec., $SD = 11.35$ vs. HC $M = 50.33$ sec., $SD = 31.43$). Using a similar paradigm (i.e., picture description task) McNamara, Obler, Au, Durso, and Albert (1992) also showed that individuals with PD produced significantly fewer words vs. aged matched healthy participants (PD $M = 57.8$, $SD = 25.4$ vs. HC $M = 88.5$, $SD = 64.2$) and vs. a group of participants with Alzheimer's disease ($M = 102.7$, $SD = 78.2$).

The finding of reduced verbal output in persons with PD is not consistent in the literature. Murray (2000), using two picture description tasks, reported no significant differences in total words or total utterances produced between individuals with PD ($M = 178.6$, $SD = 58.46$) vs. healthy controls ($M = 178.9$, $SD = 39.7$). Similarly, Ash et al. (2011) and Ash et al. (2012a) using narrative generation (i.e., story generation) via a wordless picture book (*Frog, Where Are You?*) failed to find significant differences in total words for individuals with PD vs. controls. While total spoken output measured in units of words (i.e., total words) or total duration is informative, within the study of discourse it is argued that reporting productivity in a measure grounded temporally (i.e., words per minute) is a more stable and reliable measure of productivity (Brookshire & Nicholas, 1994b).

The efficiency of verbal output measured in words per minute (WPM) also has been described in the PD literature. In a series of two studies Ash and colleagues did not find significant differences in WPM using a narrative generation task between individuals with PD without dementia vs. controls (Ash et al., 2011; Ash et al., 2012a). In contrast, these authors reported significant differences in WPM for a group of participants with Lewy Body dementia (i.e., LBD, a subgroup of individuals with either PDD or dementia with Lewy bodies [DLB]) suggesting that productivity is compromised as a function of cognitive impairment. Moreover, Illes, Metter, Hanson and Iritani (1988) and Illes (1989) using discourse samples elicited from personal narratives showed that individuals with PD were less efficient, producing fewer WPM vs. controls (Illes, Metter, Hanson, & Iritani, 1988; Illes, 1989).

Interestingly, when adjusted verbal rates (i.e., WPM after extraction of unfilled pauses > 2 seconds) of individuals with PD were compared to those with Huntington's disease, Alzheimer's dementia, and controls, no significant effects of group were reported (Illes, 1989). Illes (1989) concluded that the presence of verbal disruptions in the form of unfilled pauses > 2 seconds were important sources of reduced verbal efficiency in the discourse of individuals with PD. This finding is consistent with Illes et al. (1988) who also reported that adjusted verbal rates (i.e., pauses > 2 seconds removed), but not unadjusted verbal rates, were equivalent between individuals with PD ($M = 169.10$; $SD =$

54.20) and controls ($M = 154.50$; $SD = 25.50$). Ash et al. (2012a) also reported that significant differences in WPM between LBD and healthy controls were obviated once unfilled pauses > 2 seconds duration were removed from the WPM calculation.

Collectively, the findings among these studies indicate that unfilled pauses vs. speech production rates impair productivity in PD discourse. Discussed in more detail later in this section, the presence of unfilled pauses (i.e., > 2 seconds) contributes to verbal fragmentation of discourse in individuals with PD. While beyond the scope of the current review researchers have demonstrated that these unfilled pauses become even more problematic during interactional discourse tasks (i.e., conversation) (Griffiths, Barnes, Britten, & Wilkinson, 2012; McNamara & Durso, 2003; Rousseaux, Seve, Vallet, Pasquier, & Mackowiak-Cordoliani, 2010). These findings underscore the importance of adopting consistent and stable measures of discourse productivity (i.e., WPM) that consider the presence of such pauses as a potentially unique contributor to discourse impairments in persons with PD.

Taken as a whole, these findings suggest that productivity of verbal output is highly variable in PD. While discourse task may be a contributing factor, conflicting results of productivity impairments are evident even for studies using similar tasks (i.e., picture description or personal narratives) but different stimuli (e.g., different pictures used for elicitation of language, personal narratives of different topics,). This suggests, as have others, that even subtle task differences such as the cognitive demands of a specific topic or the amount and type of content of a pictured scene may interact with measures of productivity (McNeil et al., 2007). This is particularly important given the work of Ash and colleagues relative to the interaction of WPM and parkinsonism with dementia (Ash et al., 2011; Ash et al., 2012a). Productivity measures are particularly vulnerable to the effects of discourse sampling technique such that larger samples yield more robust data (Brookshire & Nicholas, 1994b). These findings underscore the value of collecting robust discourse samples for analyses that encompass the breadth of discourse performance across different tasks.

Collectively, the neuropsychological literature exploring cognition and language in PD suggests that factors such as global cognitive decline, disease progression measured with

the modified Hoehn and Yahr Scale (Goetz et al., 2004), and motor severity measured with the Unified Parkinson Disease Rating Scale score (UPDRS) (Fahn, Elton, & Committee, 1987) can be important sources of variability in discourse performance, particularly for productivity. However, a global challenge in the monologic discourse literature is that few studies report sufficiently comprehensive or consistent descriptions of PD participants to facilitate comparison of results with consideration of these potentially important variables. Murray (2000) reported that global cognitive status (i.e., scores on the Mattis Dementia Rating Scale) did not correlate with total words produced. However, she did report a significant and positive relationship between speech intelligibility (i.e., measured using the Assessment of Intelligibility of Dysarthric Speech (Yorkston, Beukelman, & Traynor, 1984)) and discourse productivity measures (i.e., higher speech intelligibility corresponded to greater number of total words). Unfortunately, Murray (2000) did not report measures of global motor severity specific to PD. Moreover, while Ash et al. (2012a) reported UPDRS scores (i.e., a global measure of motor severity in PD) in relation to productivity markers in discourse, they did not report data relative to speech intelligibility.

There remains a large degree of ambiguity in the findings relative to motor severity and discourse performance. For example, experienced speech-language pathologists using perceptual judgments of the presence of motor speech/voice impairments in PD (e.g., voice, articulation) judged a higher prevalence of motor speech impairments in individuals with more severe motor symptoms as measured by UPDRS scores and disease duration vs. those with less severe motor symptoms (Sapir et al., 2001). Yet, Murray (2000) reported no relationship between disease duration and total words and Ash et al. (2012a) reported no relationship between the UPDRS and WPM. Collectively, these findings from a limited number of studies suggest that while more severe motor speech impairments are associated with more severe global motor impairments and longer disease duration in PD (Sapir et al., 2001), relative to discourse productivity only speech intelligibility correlates with performance. This dissociation between general motor severity and speech intelligibility relative to discourse performance may indicate that specifically the interactions among motor speech planning/execution, cognition, and language affect discourse productivity in PD. This explanation is in keeping with the

majority of authors from across these studies who concluded that the reduced language productivity, when present in PD, is not accounted for solely by bradykinetic (i.e., slowed) motor speech movements and that cognitive and linguistic changes (e.g., reduced processing time, language formulation deficits) likely contribute to impairments of discourse productivity (Illes et al., 1988; Illes, 1989; McNamara et al., 1992). This interpretation is consistent with Murray (2000) who reported that performance on a general measure of language ability (i.e., Aphasia Diagnostic Profiles (Helm-Estabrooks, 1992)) also significantly and positively correlated with the number of total words produced by participants with PD. The assumption that both cognitive-linguistic and motor abilities interact with productivity of spoken discourse in PD also is consistent with Ash et al. (2012a) who reported that a composite measure of executive function, grammatical comprehension, and word list recall correlated significantly with WPM. The findings of Ash et al. suggest that WPM increases with better executive function performance, with increased grammatical comprehension, and with increased word list recall (i.e., verbal memory).

Overall, the findings from the current PD discourse literature highlight the relationship between both motor speech and cognitive-linguistic performance and spoken discourse productivity in PD. This is of particular interest from a rehabilitation perspective given that the pharmacological treatments for PD have resulted generally in minimal to no improvements in speech intelligibility (for a review see Pinto et al., 2004 and Sapir, Ramig & Fox, 2008). Moreover, levodopa has variable and differential effects on cognition; improving functions mediated by the dorsal striatum and worsening functions mediated by the ventral striatum (MacDonald et al., 2011). However, only a limited number of studies have considered systematically these variables relative to discourse performance in PD and those that have used a variety of measures that are not directly comparable across studies (Ash et al., 2011; Ash et al., 2012a; Murray, 2000; Murray & Lenz, 2001). Consequently, drawing confident conclusions using cross-study comparisons is challenging given this methodological ambiguity. Rigorous experimental control over the discourse tasks used to collect samples across studies as well as improved reporting of salient descriptive variables that can influence task performance (e.g., cognition, motor severity, disease staging, speech intelligibility etc.) would improve

our understanding of the spoken language productivity challenges observed in the discourse of individuals with PD and which are reported as problematic by caregivers and communication partners of persons with PD (Roberts, Finger, Jenkins, Eaton, & Orange, 2014; Whitworth, Lesser, & McKeith, 1999).

Syntax. Measures of syntactic complexity, accuracy, and diversity appear frequently in the literature on discourse in PD. Utterance length is one measure used by discourse scholars to evaluate syntactic complexity with the assumption that longer utterances are more linguistically and syntactically complex. Interestingly, studies have reported generally that the mean length of utterance (MLU) measured in words is not affected significantly for individuals with PD. Ash et al. (2011) found that participants with PD presented with equivalent MLU performance vs. controls during narrative generation tasks. Similarly Murray and Lenz (2001) reported that the differences between PD vs. controls for MLU in personal narratives (i.e., topic directed interview) only approached significance ($p = 0.08$, $\alpha = .05$). Moreover, Murray (2000) found equivalent performances for MLU between PD vs. controls using a picture description task. Interestingly, Huber and Darling (2011), using syllables vs. morphemes as the metric for utterance length, reported that individuals with PD produced significantly shorter utterances vs. controls. Their work suggests that syllable level vs. morpheme level analyses measure different constructs in PD. Reduced syllables/utterance may reflect a tendency to use shorter words and not less complex grammar, per se. However, certainly shorter, and potentially less grammatically complex utterances may also contain fewer syllables. Alternatively, these data also indicate that the length of an utterance varies as a function of discourse task given that the expository discourse task used by Huber and Darling (2011) posed different semantic, memory, and syntactic demands vs. the discourse tasks reported by Ash, et al. (2011), Murray (2000), and Murray and Lenz (2001).

Task complexity alone may not fully explain the variability in utterance length observed in PD. Holtgraves, McNamara, Cappaert and Durso (2010) studied the effects of motor asymmetry in PD on discourse performance using a semi-structured personal narrative task. By definition, PD is an asymmetrical motor disease such that one side of the body

is generally more affected than the other (See Table 1). Holtgraves et al. reported that individuals with PD with greater left side motor symptoms had significantly lower MLUs compared to those with greater right side motor symptoms. These findings suggest that in addition to task complexity side of motor symptom presentation may affect some discourse performance variables in PD. However, one conceptual issue with these MLU data collectively is that the linear relationship between MLU and syntax complexity weakens as the language system matures and as such MLU may be a more valuable tool for understanding changes in syntax/language complexity in children with developing language systems (Scarborough, Rescorla, Tager-Flusberg, Fowler, & Sudhalter, 1991). Therefore the failure to find significant differences in MLU between PD vs. controls may reflect an issue with the sensitivity of the measurement tool vs. a reflection of equivalent syntax production abilities in discourse. At best, MLU should be considered in conjunction with other potentially more sensitive measures of syntax complexity.

Fortunately, researchers have conducted other analyses that are more sensitive to syntax impairments, across multiple discourse genres. Murray and Lenz (2001) reported no significant differences between PD vs. controls during topic directed interviews for proportion of grammatical sentences, proportion of complex sentences, number of embedded clauses per utterance, or proportion of verbs inflected. However, they did find robust correlations for the PD group between a global measure of aphasia and MLU and also between the aphasia measure and the proportion of complex sentences produced. Furthermore, Murray and Lenz (2001) reported robust correlations among performances on the Dementia Rating Scale (DRS) (Mattis, 1988) and MLU, proportion of complex sentences, and embedded clauses per utterance. These findings are consistent with the literature on syntax processing and production in the neuropsychology literature, which also report significant relationships between cognition and syntax abilities in PD.

Murray and Lenz (2001) also reported statistically significant correlations for the PD group between speech intelligibility and MLU, proportion of complex utterances and number of embedded clauses per utterance. Their findings suggest that as speech intelligibility declines, individuals with PD simplify the structure of their utterances perhaps in an attempt to conserve linguistic or motor resources. However, alternatively

these data also suggest that there are commonly disrupted processes in PD that affect both language and speech production. While the exact reason for the relationship between syntax and motor speech complexity remains unresolved, these findings underscore the importance of discourse as one method for exploring the integration of cognition, language, and motor speech in syntax production.

In contrast to Murray and Lenz's (2001) syntax findings using a personal narrative discourse task, Murray (2000) did report significant differences in proportion of grammatical sentences produced using a picture description task for PD ($M = .73$; $SD = .12$) vs. controls ($M = .92$; $SD = .08$). There were no significant group differences for other measures of syntax complexity (i.e., proportion of simple sentences or embedded clauses per utterance). Collectively, the findings from Murray (2000) and Murray and Lenz (2001) indicate that the intrinsic demands of a particular discourse task (e.g., the higher semantic and lexical constraints of a picture description task vs. a personal narrative task) can affect syntax production abilities, specifically the ability to produce grammatically correct and complete sentences. Consistent with Murray and Lenz's (2001) data from personal narratives, Murray (2000) also reported significant correlations among measures of language, cognition, and discourse performance using picture description tasks. Taken as a whole, these findings show that length of utterance shortens and syntax becomes less complex as a function of more impaired performance on cognitive and language measures (Murray, 2000; Murray & Lenz, 2001). Interestingly, using a picture description task (Murray, 2000) reported no relationship between measures of speech intelligibility and MLU or other measures of grammatical accuracy/complexity. This contrasts with findings from personal narrative tasks (Murray and Lenz, 2001) and may indicate that the two tasks place differing demands on the motor speech system. Task specific effects on motor speech output in PD have been reported previously (Kempler & Van Lancker, 2002). However, of interest, Murray (2000) reported very robust positive correlations between the lexical retrieval subtest of the Aphasia Diagnostic Profiles (Helm-Estabrooks, 1992) and MLU suggesting that for the picture description task word retrieval abilities were associated with MLU vs. the

motor speech association that was observed in Murray and Lenz's study using a personal narrative task.

In concert, these two findings suggest that the added lexical retrieval demands of the picture task interacted with MLU, whereas for the personal narrative task, motor speech demands interacted with MLU. Collectively, these results underscore the complexity of syntax production within contextual communication tasks and highlight the advantage, as Nicholas and Brookshire (1994b) suggested over twenty years ago and more recently McNeil et al. (2007) and Marini, Boewe, Caltagirone, and Carlomagno (2005) have suggested, of studying discourse performance across multiple genres vs. solely within a single genre.

Illes et al. (1988) and Illes (1989) used a syntactic scoring system based on complexity of clausal structures within sentences (i.e., embedded vs. non-embedded clauses) to analyze syntax complexity for personal narratives in persons with PD. Illes et al. (1988) reported that their participants with PD did not differ significantly from controls in the complexity of syntax structures produced during spoken discourse. However, Illes et al. reported that within the PD group individuals with mild PD (i.e., determined by the Webster 30-Point scale of Parkinsonian Disability the appropriate standard for the time of her publication (Webster, 1979)) produced sentences with a significantly higher degree of syntactic complexity (i.e., more embedded clauses) vs. those with moderate PD severity (Illes et al., 1988). In addition, the syntax complexity scores correlated significantly and robustly with the Webster scale. This latter finding was consistent with Illes follow-up study (Illes, 1989). With increased disease severity, the PD participants from Illes et al. (1988) and Illes (1989) produced sentences that were less syntactically complex (i.e., contained fewer embedded clauses). Additionally, Illes et al. (1988) observed that sentences produced by individuals with more severe PD vs. those with less severe PD were characterized by strings of open class phrases organized in a list-like fashion (vs. embedded clauses) resulting in longer sentences with reduced syntactic complexity. Below is an excerpt from Illes et al. (1988), illustrating this discourse feature in an individual with PD. The speaker produced five open class phrases (e.g., content word

phrases) in a list-like fashion all modifying the main clause “I worked”. Each open class phrase is underlined:

“ I worked for thirty-two years for the [the] Department of Water and Power, first as a [a] /m/ [/m - m/] mechanic, then a [a] lead man, then finally as a /f/ [/f/] foreman,” (Illes et al. 1988, p. 156)

These findings in context with those of Murray (2000) suggest spoken discourse in PD is marked by reduced syntax complexity and grammatical accuracy that, while influenced by motor speech changes, is not the sole result of motor speech changes. Moreover, these findings suggest that the mechanisms utilized by individuals with PD to create longer sentences differ from controls and typically involve the linking together of multiple independent clauses vs. use of embedded or dependent clauses. Likely, as others have suggested, these findings result from interactions among cognitive, language, and motor changes in PD and/or the common underlying neurological changes serving all three process.

The results from Ash et al. (2011) and Ash et al. (2012a), in which she and her colleagues used a narrative generation task, are in general agreement with those presented by the Illes studies (Illes et al., 1988; Illes, 1989) and Murray studies (Murray, 2000; Murray & Lenz, 2001) suggesting that some, but not all, individuals with PD present with reduced syntax complexity/grammatical accuracy in spoken discourse. In Ash et al. (2011), the percentage of well-formed utterances (i.e., utterances containing a subject and a verb clause, free of grammatical errors, and correct to the story) and the percent of utterances with complex structures (i.e., dependent clauses and phrasal adjuncts) differed from controls only for participants with PD or parkinsonism and dementia. Ash et al. reported similar findings in their 2012 study using a composite measure for syntax complexity and grammar accuracy vs. evaluating the constructs individually (Ash et al., 2012b). Interestingly, these findings occurred in the presence of equivalent UPDRS-III scores between the dementia-affected and unaffected participants indicating that the observed findings result from cognitive differences and not motor differences between the PD groups. While these findings suggest that syntax/grammar production impairments are

present in the discourse of individuals with PD and dementia and not for individuals with PD without dementia what remains unclear is whether or not syntax complexity is impaired in the presence of more subtle cognitive changes such as those occurring in PD mild cognitive impairment (PD-MCI).

Collectively, this group of studies reporting syntax production in spoken discourse tasks among persons with PD presents what on the surface appear to be conflicting findings with some studies showing PD-related impairments in syntax production vs. others reporting differences exclusively in the context of concomitant dementia. However, the challenge lies in the attempt to interpret these data as a collective body of literature. It is likely that discourse task differences across studies contributed to these conflicting findings. Murray (2000) used a picture description task whereas the Ash studies (2011 and 2012a) used narrative generation and both Murray and Lenz (2001) and the Illes studies (1988 and 1989) used personal narratives. The increased semantic and lexical constraints of the picture description task in Murray (2000) may have posed word retrieval challenges for the PD group that interacted with the ability to access and to produce more complex syntax structures. Additionally the nature of the stimuli used in the picture description task may have encouraged individuals with PD to produce more list-like descriptions of the events in the picture similar to the language produced in the more severe PD group in Illes, et al. (1988).

The existing body of literature suggests that cognitive status and disease severity interact with the intrinsic demands of the discourse task to affect the complexity of syntax and grammatical accuracy in spoken language production in PD. However, Illes et al. (1988) and Illes (1989) did not use any formal measure of global cognitive function to assess their participants with PD (i.e., they used self-report only). In contrast, the Murray studies (Murray, 2000; Murray & Lenz, 2001) used the Dementia Rating Scale (DRS), an earlier version of a recommended global cognitive measure for detecting PD-MCI and PDD (Marras et al., 2013). This strengthens the Murray and Murray and Lenz studies not only in their participant selection criteria but also in the strong data they provide reporting correlations between measures of global cognition and measures of syntax complexity within their PD group. The Ash et al. (2011) and Ash et al. (2012a) studies,

in addition to published diagnostic criteria for PDD/DLB applied by a neurologist, used the Mini Mental Status Examination (MMSE) (Folstein, Folstein, & McHugh, 1975) (with a cut-off score of 24) to classify individuals with PD into dementia vs. non-dementia groups. The Dementia Rating Scale - 2 (Mattis, 2001) was used as a secondary measure of global cognition but only for participants who were felt to exhibit dementia despite scoring above 24 on the MMSE (no participant numbers for whom this was required were reported by the authors). A recent study by Marras et al. (2013), reported that using MMSE score of 24 (i.e., the cut-off score for normal vs. dementia used in Ash et al. 2011; Ash et al., 2012a) the sensitivity for detecting PD-MCI was .152 and specificity was .989. This could suggest some ambiguity in the cognitive assignment of individuals in the Ash et al. (2011) and Ash et al. (2012a) studies. Another challenge is in understanding the relationship between disease/motor severity and syntax and grammar performances in spoken discourse. A challenge with accepting the face value of Ash et al.'s (2011 and 2012a) findings that motor severity using the UPDRS-III was equivalent between the dementia-affected and unaffected PD groups is that motor severity may differ between individuals with PDD and those with DLB. Symptoms of postural instability, rigidity, and gait disruption, which inflate scores on the UPDRS-III are more prominent in individuals with PDD vs. DLB (Burn et al., 2003). It is possible that the DLB UPDRS-III scores in the mixed dementia participant group lowered the mean score obscuring potentially higher UPDRS-III scores for individuals with PDD and thus obscuring potential effects of motor severity on the findings. These methodological concerns reinforce the importance, particularly when using syntax analyses in discourse, of collecting robust participant descriptive data specific to PD and to carefully control for the presence of cognitive changes that meet the criteria for dementia vs. those subtle changes that may be typical of the disease.

Lexical diversity. The choice, accuracy, and diversity of words produced are important measures of monologic discourse performance in adults with progressive neurological disorders affecting cognition and language. In PD, lexical skills interact with and overlap with cognitive processes, syntax, semantic access and errors, coherence and cohesion processes, accuracy of conveying information, verbal disruptions, and

motor components of speech programming. Research in PD suggests that word retrieval impairments and ‘tip of the tongue’ phenomenon are considerable (Bertella et al., 2002; Boulenger et al., 2008; Cotelli et al., 2007; Matison, Mayeux, Rosen, & Fahn, 1982; Peran et al., 2009; Rodriguez-Ferreiro et al., 2009). Consequently, studying lexical diversity in PD within discourse paradigms is important because such analyses are likely to reveal word retrieval challenges that are not apparent in standardized language testing.

Calculating the proportion of open class and closed class words is one method used by discourse scholars for analyzing lexical diversity in spontaneous language. Ash et al. (2011) and Ash et al. (2012a), using a narrative generation task, did not find significant differences in the proportion of open class words (i.e., content carrying word classes such as nouns, verbs, adverbs, and adjectives) produced in the discourse samples of individuals with PD vs. controls. All of the participant groups (i.e., PD, LBD and controls) produced approximately 40% open class words. Similarly, Murray and Lenz (2001) reported no significant differences between individuals with PD vs. controls in the proportion of closed class words (i.e., non-content conveying words: prepositions, determiners, conjunctions, pronouns) with both groups producing approximately 50% closed class words (Murray & Lenz, 2001).

Surprisingly, there is a relative paucity of data on lexical diversity in discourse production in PD compared to other areas of discourse analyses reported in the literature. Not surprisingly, the data suggest that discourse genre may play an important role in the lexical diversity profiles demonstrated in PD. In fact, recent work has demonstrated that type-token measures of lexical diversity even those based on computational and statistical methods (e.g., *D*) are particularly vulnerable to discourse task differences in both healthy adults and in adults with aphasia, which may affect the reliability of these measures in cross-genre discourse studies (Fergadiotis & Wright, 2011; Fergadiotis, Wright, & Capilouto, 2011). However, the specific effect of task on lexical diversity in PD has yet to be explored. Discourse genres such as personal narratives that focus on sharing a personal recount of information or life events may bias the types of words used via the types of syntax structures that are produced and increased tangentiality (Marini et al., 2005). In contrast to personal narrative genres, discourse tasks with a high degree of

semantic and lexical constraint (e.g., picture description, story retelling), which may reduce tangentiality may increase the proportion of open class words. Yet, these types of highly constrained tasks have rarely been explored in PD spoken discourse studies. This is important given that in individuals with aphasia, the use of picture supported discourse elicitation stimuli have been shown to result paradoxically in samples with increased lexical diversity vs. tasks that do not have picture support (Fergadiotis and Wright, 2011; Fergadiotis et al., 2011). This is a finding which Fergadiotis and Wright (2011) and Fergadiotis et al. (2011) hypothesized was the result of the additional scaffolding for word retrieval that was provided by the picture which may have compensated for the increased semantic and lexical constraints typically assumed to be present in picture description tasks.

The existing literature in PD discourse indicates that individuals with PD use a balance of open vs. closed class words that is similar to that of individuals without PD. In other words, their language is not telegraphic (i.e., reduced use of closed class words) as in Broca's aphasia. Task differences but not disease status per se appears to interact with measures of lexical diversity in PD. Collectively, these studies help to inform the need for developing systematic and consistent discourse elicitation methods and analyses in order to optimize the utility of lexical diversity both as a research and clinical tool in PD.

Verbal fragmentation analyses. Verbal fragmentation analyses quantify the frequency and type of verbal disruptions (i.e., disruption in the fluent and fluid output of information) within spoken discourse. Verbal disruptions can result from a host of cognitive, linguistic, and motor speech sources. For example, verbal disruptions in monologic discourse, such as unfilled pauses, can result from lags in information processing resulting in response delays. Additionally, unfilled pauses can result from delays in selecting and/or retrieving words from the lexicon or be the mechanism to hold one's turn in speaking even within a monologic task. Pauses that disrupt the flow of spoken output also may result from delays in selecting, retrieving, and executing the motor programs for a particular word, utterance, or sequence of utterances. Repetitions of sounds, syllables or words also can reflect cognitive, linguistic or motor speech challenges in planning and execution at phonological or lexical levels. Verbal

disruptions such as reformulations or revisions can reveal deficits in word retrieval or syntax selection and/or execution.

The very integrative and interdependent cognitive-linguistic and motor processes required for the production of continuous spontaneous spoken language is such that isolating the specific source of verbal disruptions can prove challenging within the context of discourse vs. well-controlled single word production tasks, and is particularly magnified in studies of persons with PD. Still the study of verbal disruptions in discourse is incredibly valuable not only because of their sensitivity for revealing challenges in these integrative processes of spoken language but also because of the role such disruptions play in how the listener ultimately perceives and processes both the speaker and their spoken language (Merlo & Mansur, 2004; Roberts, Meltzer, & Wilding, 2009). As such, the study of verbal disruptions in PD is probably the most ubiquitous metric applied across discourse studies.

Huber and Darling (2011) conducted linguistic and motor speech analyses on expository discourse and on oral reading samples of participants with PD and healthy adult controls. They reported both quantitative and qualitative differences between individuals with PD vs. controls. While both groups produced significantly more formulation errors (i.e., abandoned utterances, revised phrases, phrase repetitions) in the discourse task vs. oral reading, individuals with PD produced significantly more formulation errors vs. controls for both tasks. Individuals with PD produced significantly fewer filled pauses (e.g., pauses denoted with a verbal filler such as “um” or “uh”) vs. controls during the discourse task whereas during the oral reading task, the proportion of filled pauses was equivalent between PD and controls. In contrast, Huber and Darling (2011) reported that the proportion of disfluencies (e.g., repetitions of sounds, single words, or syllables) between the groups during expository discourse was equivalent. These findings suggest that during spoken discourse participants with PD produced more pauses and revisions (i.e., words and phrases) but not more disfluencies (i.e., repeated sounds, words) vs. controls (Huber & Darling, 2011). The finding that verbal disruptions in the PD participants occurred primarily at the phrase and utterance level vs. the sound or word

level suggest, although not conclusively, that these disruptions are reflective more of cognitive-linguistic challenges vs. articulation or voice production challenges.

Similarly, using procedural narratives, Illes et al. (1988) reported that the number of unfilled pauses > 2 seconds duration/minute was significantly higher for individuals with PD. Importantly, Illes et al. showed that the presence of unfilled pauses > 2 seconds (i.e., silent hesitations) was one of four variables reported to separate individuals with PD vs. controls. Moreover, these researchers reported that the unfilled pauses occurred most commonly at the beginning of sentences and at the boundaries between mandatory and optional phrases within the sentence. Consistent with Huber and Darling (2011), Illes et al. (1988) reported that the number of filled pauses/minute (e.g., “uh”, “uhm”) did not differ between PD vs. controls. These findings suggest that the presence of unfilled pauses may be a salient feature of verbal disruptions in the discourse of individuals with PD. While disambiguating the exact source of these unfilled pauses (i.e., cognitive-linguistic or motor) may be challenging the location of unfilled pauses suggests an interaction between this form of verbal disruption and both language and motor formulation in PD.

Illes (1989) investigated the profile of verbal disruptions during a personal narrative discourse task in a group of individuals with PD who were divided into mild and moderate disease severity groups. Those with moderate severity PD produced significantly more unfilled pauses vs. controls. Interestingly, relative to the proportion of unfilled pauses > 2 seconds, individuals in the moderate severity PD group performed similarly to a comparator group of participants with moderate severity Alzheimer’s dementia, a non-motor disease of cognitive impairment whose verbal disruptions largely reflect language planning deficits. PD disease severity also may play a role in the prevalence of unfilled pauses. For example, while unfilled pauses were an issue for the moderate PD group, Illes (1989) found no difference for the proportion of unfilled pauses between the mild-disease PD group vs. controls. This finding is in agreement with Jaywant and Pell (2010) who reported no significant differences in either mean duration of pause or mean percent pause time on a picture description task for individuals with mild PD severity vs. controls.

Overall, disease severity also appears to play a role in the prevalence of filled pauses as a form of verbal disruption. Illes and her colleagues (1988) reported that the occurrence of filled pauses (i.e., pauses containing verbal behaviors used to fill time such as ‘uh’ and ‘uhm’) separated the mild vs. the moderate PD participants, with the mild PD participants, like controls, producing significantly more filled pauses vs. those in the moderate severity group. Elaborating on the quantity and quality of verbal disruptions in PD, Illes (1989) reported that the proportion of interjections (i.e., exclamatory phrases in the stream of language production) and modalizations (i.e., comments made by the speaker about the task or about their performance) also were significantly higher for controls (i.e., more than twice that of PD). Interestingly, and consistent with other published studies, the proportion of word and sound level verbal disruptions (i.e., semantic paraphasias, neologisms, and phonemic paraphasias) was not significantly different between the PD and control participants (Illes, 1989).

Ash et al. (2011) using a narrative generation paradigm expanded the description of verbal disruptions in PD vs. controls by quantifying both the proportion of phonemic errors and phonetic errors produced. Ash et al. (2010) defined phonemic errors as speech production errors that result in a word containing sound(s) that are “well-formed” for the language but not “intended or anticipated by the listener” such as “coming out of the gar” for ‘coming out of the jar’ (p. 14). Phonetic errors were defined as speech production errors that result in a word containing a “sequence of sounds” that are “not possible” in the speaker’s spoken language such as reduced articulation force/pressure on a target consonant such as softening the /g/ in “dog” resulting in a sound distortion that does not represent an English sound) (Ash et al., 2010, p. 14). Ash et al. (2011) reported no significant difference between individuals with PD vs. controls in either phonetic or phonemic errors. However, in the context of dementia (i.e., individuals with PDD/DLB) a significant difference vs. controls emerged, with individuals with PDD and those with parkinsonism and dementia (i.e., DLB) producing significantly higher rates of phonetic and phonemic errors than either the PD or control groups.

Using the same discourse elicitation paradigm, Ash et al. (2012a) calculated a composite variable they termed “articulation errors” comprised of: 1) proportion of phonetic or

phonemic errors, 2) proportion of false starts (i.e., incomplete words subsequently abandoned), 3) proportion of filled pauses (i.e., defined as hesitation markers and editing breaks, and 4) proportion of disfluent words (i.e., self-corrected words, word revisions). These disruptions were calculated as a proportion of ‘articulation errors’ per 100 words. Individuals with PDD/DLB produced significantly more errors ($M = 17.3$; $SD = 11.9$) vs. controls ($M = 5.9$, $SD = 6.0$). However, no differences were reported between individuals with PD ($M = 6.4$, $SD = 5.0$) vs. controls. Ash et al. also explored the presences of pauses as a measure of verbal disruption. In Ash et al. (2012a), pauses exceeding 2 seconds in duration were calculated as a proportion of the total narrative duration. The proportion of verbal disruption pauses was significantly higher for both the DLB/PDD subgroup ($M = 37.3$, $SD = 17$) and PD subgroup ($M = 13.1$, $SD = 14.3$) vs. controls ($M = 4.6$, $SD = 6.0$) (Ash et al. 2012a).

The studies by Ash and her colleagues (2011 and 2012a), like other studies, suggest that unfilled pauses are a substantial source of verbal disruption in PD even in the absence of dementia. However, using a more detailed analysis of word level errors in discourse, Ash et al. (2011) suggested that the presence of more motor-based verbal disruptions becomes problematic only in the presence of increased cognitive impairment (Ash et al., 2011). While it remains important to consider their assertions relative to the selection of global assessments of cognition used across discourse studies, the findings of Ash et al. (2011) and Ash et al. (2012a) are particularly intriguing relative to the interaction between cognitive impairment and motor-based verbal disruptions. These findings underscore the important influence of cognitive processes on motor abilities within spontaneous language production and the need to consider this in the design of and interpretation of discourse impairments in PD.

Additional sources of verbal disruptions in discourse production can occur as a speaker attempts to correct an error in speech or language production. These repair attempts can result in verbal disruptions occurring at the sound, syllable, word, or phrase level such as a) repetitions and b) revisions or reformulations. McNamara et al. (1992) explored the ability of individuals with PD, AD and controls to self-monitor and correct language errors during spontaneous language using a picture description task. Individuals with PD

produced significantly more undetected language errors (75%) vs. age-matched control participants (18%). In fact, the percentage of undetected errors in a cohort of individuals with PD without dementia (criteria for determining was not specified in the study method) was consistent with the undetected error rates of the AD group (76%). These findings help inform an interpretation of the data from Illes (1989) who reported a high prevalence of self-correction verbal disruptions in AD and HD, but minimal use of self-correction verbal behaviours (i.e., reformulations, revisions) in the PD group. Taken as a whole, one possible interpretation of the Illes findings could be that fewer episodes of verbal disruptions typically associated with self-corrections (e.g., repetitions, revisions) are observed in PD because these individuals are less able to either detect errors or develop on-line strategies for correcting these errors. McNamara et al. (1992) also explored the types of repair strategies applied within discourse. Only 14% of language errors were corrected using a reformulation strategy (i.e., multiword correction where a new lexical constituent is added) in the PD group vs. controls who used reformulation strategies to repair detected errors 35% of the time (statistically significant). Similarly, individuals with PD used significantly fewer lemma repair strategies (i.e., correction where an error is replaced by a single word) vs. controls with 11% of total errors and 47% of total errors, respectively.

Collectively, the studies reviewed suggest that understanding verbal fragmentation in the spontaneous spoken language of individuals with PD is complex. While it is true that the inherent nature of spontaneous language presents challenges in disambiguating fully and isolating a target source (i.e., cognitive-linguistic or motoric) for specific verbal disruptions, the literature to date relative to the types of verbal disruptions and the contexts in which verbal disruptions occur in PD suggest that these behaviours result from the integration of motor and cognitive processes within the context of language production. While the characterization of verbal disruptions may differ as a function of language task, motor severity, and cognitive status, the current literature suggests that the presence of verbal disruption impairments are ubiquitous in PD and therefore an important metric for the study of discourse performance in PD.

Macrostructural analyses. The following is a discussion of the nascent studies and their findings from the spoken monologic discourse literature in PD relative to thematic unity and information-level analyses (i.e., coherence, cohesion, main idea analyses, correct information unit analyses)

Local and global coherences. Few studies have explored the discourse of individuals with PD from the perspective of creating and maintaining unifying themes within narratives from the perspectives of local and global coherence. Coelho (2007) defined local coherence as “the relationship of the meaning or content of an utterance to that of the preceding utterance” (p. 124). He further defined global coherence as “the relationship of the meaning or content of an utterance to the general topic of the story” (p. 124). Given the frequency of higher-level language impairments (e.g., figurative language), working-memory impairments, and executive function impairments reported in PD, it is reasonable to expect impairments in local and global coherences in their spoken discourse (Berg et al., 2003; Farag et al., 2010).

Only one published study to date has explored local and global coherences of spoken discourse in individuals with LBSD (Ash et al., 2011). The unique scoring system used by Ash and colleagues, based on the scoring of individual events in the spoken narratives, contained measures for: a) local connectedness (i.e., event presented relates to preceding material established via linguistic connecting devices), b) accurate identification and maintenance of the central theme of the story, and c) global connectedness (i.e., identifying the correct resolution of a problem in the story and linking back the resolution to the main theme and the initiating event) (Ash et al., 2011). Ash et al. reported that while individuals with PD performed worse than controls on all measures of local and global coherence, these differences failed to reach statistical significance.

Interesting, Ash et al. (2011) found that significant group differences emerged only in the presence of dementia. The LBD group (i.e., participants with PDD or DLB) was more impaired than both the PD group and controls on measures of local and global connectedness and coherence. Ash et al. (2011) reported that measures of local connectedness within the LBD group correlated robustly and negatively with motor

severity, disease staging, and verbal output productivity (words/minute) (Ash et al., 2011). In addition, they found that for participants in the LBD group, measures of local connectedness correlated robustly and positively with measures of executive function (i.e., reverse digit span, phonemic fluency, and semantic fluency). In other words, more impaired performance in maintaining local connectedness was associated with greater motor severity, more progressed disease staging, and worsening cognition. Moreover, as the ability to maintain local connectedness declined so did productivity of spoken discourse as measured by WPM. The ability to identify accurately and to maintain the central theme of the story correlated robustly and positively with verbal output productivity, phonemic fluency, semantic fluency and negatively with Stroop Color Word testing for individuals in the DLB group. These findings suggest that the ability to create and to maintain a central theme of a story was related highly to cognitive performance. Measures of global connectedness correlated significantly and negatively only to disease staging (i.e., Hoehn and Yahr) suggesting that for individuals with PD and dementia the ability to identify resolutions to a problem in a narrative and to relate that to the initiating event (i.e., source of the problem) declines with more advanced disease.

The findings of Ash et al. (2011) suggest that while individuals with PD without dementia performed worse than controls on measures of local and global coherence, these differences reached statistical significance only for those individuals with dementia. This finding is in keeping with the larger body of literature in discourse relative to measures of connectedness and cohesion, suggesting that these processes are sensitive to changes in cognitive ability, specifically measures of global coherence (Coelho, 2007; Wright, Koutsoftas, Capilouto, & Fergadiotis, 2014). There is emerging evidence to suggest that measures of local connectedness and global coherence in the discourse of individuals with PD are disrupted only in the presence of dementia; however, these studies require replication and expanding. This area of work in PD is certainly worth further exploration in light of newly available diagnostic criteria for PD-MCI (Litvan et al., 2012) to determine at what level of cognitive disruption in PD (PD-MCI vs. PDD) do performance differences emerge in connectedness and cohesion devices within discourse. This is of particular importance given that these areas of discourse performance can create substantial communication challenges when disrupted.

Informativeness. The degree of informativeness in spontaneous language often is measured along the dimensions of the amount, quality, and efficiency of information exchanged between the speaker and the listener (Shadden, 1998b). Given that one of the primary goals of communication is to convey information, measures of informativeness are of critical importance in the understanding of discourse abilities (Shadden, 1998b). There are multiple methods for measuring informativeness within the context of discourse. While subtle variations in analyses systems exist, ‘main idea’ or ‘content unit’ analyses typically evaluate discourse samples for the quantification of groupings of words (i.e., words, clauses or sentences) that reflect one or several main themes expressed in a particular stimulus (e.g., picture). They are typically scored using a binary scale as either present or absent from the discourse sample (Capilouto, Wright, & Wagovich, 2005). The lists of themes are developed a priori, are stimulus-specific and are based on normative studies conducted typically with healthy adults. Several researchers have validated the use of main idea analyses in discourse (Capilouto et al., 2005; Nicholas & Brookshire, 1995; Yorkston & Beukelman, 1980). A commonly used information analysis applied to discourse is ‘correct information units’ (CIUs). CIU analysis, often expressed as % CIUs or CIUs/minute, is a rule-based system for quantifying the number of words that are correct and reflect novel information relative to the stimulus (Capilouto et al., 2005; Nicholas & Brookshire, 1993). Main idea analyses and CIU analyses reflect complementary vantages of informativeness. Main idea analyses capture how groups of words relate to the overall global context of a story while CIU analysis is conducted at the lexical level and reflects the proportion of words produced in a discourse sample that are correct, novel, and informative relative to the picture stimulus (Capilouto et al., 2005; Nicholas & Brookshire, 1993). While disruptions in discourse informativeness are well documented in the literature for many progressive and acquired neurological cognitive-communication disorders such as AD, amyotrophic lateral sclerosis, acquired brain injury, Huntington’s disease, and frontotemporal dementia (Coelho, 2007; Fleming & Harris, 2008; Jensen, Chenery, & Copland, 2006; Murray, 2000; Murray & Lenz, 2001; Orange & Kertesz, 2000; Roberts-South et al., 2012; Zraick et al., 2011), few studies have explored measures of informativeness in PD.

McNamara et al. (1992) investigated the amount and the efficiency of information conveyed using a picture description task in PD with main idea analyses. Individuals with PD produced a comparable number of units of information as age-matched controls. However, they produced significantly fewer words per information unit (8.4 words) vs. controls (15.2 words). These findings suggest that while individuals with PD did not omit any of the critical themes of information, the productiveness of how these themes of information were conveyed differed significantly from controls. In another study using a picture description task, Murray (2000) found significant differences in informativeness between individuals with PD and controls using CIUs. Murray reported that individuals with relatively mild-moderate PD without dementia produced significantly lower percent CIUs ($M = 76.69$, $SD = 10.91$) vs. controls ($M = 89.73$, $SD = 5.24$). Moreover, individuals with PD produced a significantly lower percentage of informative utterances defined as complete utterances that communicate accurate and novel information relative to the picture ($M = .85$, $SD = .12$) vs. controls ($M = .96$, $SD = .04$).

Using a different discourse genre, narrative-retelling task, Bayles (1990) identified 25 possible a priori units of information in her main idea analysis. Consistent with McNamara et al. (1992), individuals with PD did not differ significantly in the amount of information produced vs. controls. However, significant differences were found when Bayles (1990) merged the participant groups and allocated them (i.e., both PD and controls) into either a high-normal cognition group or a low-normal cognition group based on scores from a standardized measure of global cognition. In her sub-analysis, the high-normal cognition group, independent of having PD, produced significantly more units of information vs. the low-normal cognition group. Bayles's findings, interpreted with caution given the use of arbitrary cut-off points to divide the high-normal vs. low-normal cognition groups, suggest that cognitive status is an important variable to consider when analyzing the amount of information produced via main idea analyses in a high declarative memory loaded discourse genre such as story retelling in PD.

The ability to generate the required number of and correct sequence of steps for a procedure or task (i.e., procedural discourse) is critical in everyday activities. Godbout and Doyon (2000) explored the performances of individuals with PD without dementia

and a control group on a series of procedural discourse tasks requiring both forward and backward generation of scripts. Spoken procedural discourse tasks assess the ability to access and to convey correctly, using spoken language, knowledge of the prototypical events and steps associated with completing familiar tasks.

Consistent with the McNamara et al. (1992) and the Bayles (1990) studies, Godbout and Doyon (2000) reported no significant differences between PD vs. controls for the total units of information produced. However, there were differences in the quality of information between the PD and control groups. Individuals with PD produced significantly more errors in sequencing the order of information required to perform the targeted procedure. In the forward condition, 12/16 individuals with PD produced sequencing errors compared to 1/16 controls. These findings suggest that individuals with PD have challenges in sequencing information in spoken procedural narratives even when the amount of information is sufficient. Godbout and Doyon (2000) suggested that the challenges faced by individuals with PD in procedural discourse result from an impaired ability to order events temporally in episodic memory.

In addition to episodic memory challenges, Godbout and Doyon's (2000) study highlights other cognitive challenges manifested by individuals with PD in conveying relevant information during procedural discourse. The PD participants in Godbout and Doyon produced perseveration errors (i.e., repeating information already provided) even in the forward condition (5/16 participants), whereas controls did not produce any perseveration errors. The authors interpreted this finding to suggest that some individuals with PD have challenges managing attention resources (i.e., set-shifting) during procedural discourse that consequently affect information content of narratives (Godbout & Doyon, 2000). The authors also reported that their participants with PD produced significantly more irrelevant units of information (e.g., information not correct to the targeted procedure). Godbout and Doyon (2000) concluded that individuals with PD have a specific impairment in accessing event knowledge or maintaining scripts for event knowledge both of which negatively impact informativeness of spontaneous language.

Family members complain of reduced informativeness or reduced information complexity in the spontaneous language of their relatives with PD (Roberts et al., 2014; Whitworth et al., 1999). Naive listeners also rate the spoken language of individuals with PD as “less detailed” or less informative (Jaywant & Pell, 2010; Whitworth et al., 1999). While these differences in informativeness may be observable by communication partners, the studies reviewed suggest that revealing these impairments in PD, within the context of structured discourse tasks, requires carefully considered and systematically selected metrics. For example, measuring informativeness using main event or content unit analyses in isolation may be less sensitive to the types of discourse deficits seen in PD. Collectively, the discourse studies in persons with PD suggest that, with the exception of discourse tasks that particularly tax memory components, it is more appropriate to detect informativeness changes using word level analysis (i.e., correct information units) vs. using larger thematic level analyses. While these data are limited, findings from discourse studies in PD present preliminary evidence to support the finding that the accuracy, amount, and efficiency of informativeness is affected in PD and that the ability to sufficiently and efficiently convey information in spontaneous language may be highly influenced by the inherent demands of the discourse task.

2.4 Background Summary

Collectively, findings from the limited literature suggest that the neurological consequences of PD result in cognitive, linguistic, and motor impairments that are uniquely revealed via performance on spoken monologic discourse tasks.

Notwithstanding the limited number of studies and diverse methodologies, there is preliminary evidence to support a profile of impairments in spoken monologic discourse in individuals with PD that includes:

- impaired productivity when measured in words per minute;
- reduced use of complex sentences including sentences with embedded clauses;
- reduced fluency and fluidness of verbal output particularly marked by unfilled pauses likely resulting from both cognitive-linguistic and motor sources; and
- reduced volume and efficiency of information conveyance particularly when measured with lexical level analyses.

The collective findings lend support for the use of discourse tasks and analyses to understand communication challenges in PD. However, our knowledge of cognitive-linguistic profiles in persons with PD would be expanded and enhanced by systematically and rigorously designed multi-level, multi-genre discourse studies that are sufficiently powered by larger sample of individuals with PD and controls.

2.5 Rationale for the Current Study and Statement of the Problem

Given the nature of our globally aging society and the importance of communication in the everyday lives of individuals, it is prudent that researchers and clinicians expand their understanding of the cognitive, language, and communication challenges occurring in those neurodegenerative diseases, such as PD, that disproportionately affect older adults. The review of the current literature demonstrates that systematically designed and carefully crafted research studies, using spoken monologic discourse tasks, can offer valuable insights into the nature of such impairments in PD from the perspective of the integration of cognitive, linguistic, and motoric components of communication.

Yet, to date, few researchers have explored comprehensively and rigorously the discourse performance of individuals with PD on such tasks. Moreover, recent advances in the prevailing knowledge of cognitive and language impairments in PD without dementia inform the design of discourse paradigms that optimize the current state of knowledge in the field. Such a study could provide a valuable foundation for informing clinical interventions and developing research protocols addressing communication challenges in LBSD (i.e., continuum of neurologically progressive disorders that includes PD, PD-MCI, PDD, and dementia with Lewy bodies).

2.6 Research Questions (RQ), Hypotheses, and Objectives

Study Objective: The objective of the current study was to create a profile of spoken discourse impairment in PD that, with acceptable sensitivity and specificity, distinguished the spoken discourse of individuals with PD from the spoken discourse of healthy older controls.

Hypothesis: Changes in cognition and language, occurring in the context of PD, affect spontaneous language in a predictable manner that can be uniquely characterized using spoken monologic discourse tasks.

RQ 1: On which measures do participants with PD differ significantly from controls using a comprehensive battery of standardized assessments of cognition, expressive language, and speech intelligibility?

RQ 2: Does performance on measures of discourse productivity, lexical use, grammar, informativeness, and verbal fragmentation differ significantly between PD vs. controls as a function of discourse task using a cross-genre sampling method and a multi-level discourse analyses paradigm?

RQ 3: For which domains (i.e., productivity, lexical use, grammar, informativeness, verbal fragmentation) and on which specific discourse measures, do participants with PD differ significantly from controls using a cross-genre sampling method and a multi-level discourse analyses paradigm?

RQ 4: To what degree does a unique profile characterizing discourse impairments in PD discriminate the spoken language of participants with PD from that of controls?

RQ 5: What is the nature of and strength of the relationship between discourse performance and markers of disease severity in PD, age, education, and global cognitive function?

Chapter 3: Method

A cross-genre, multi-level discourse analyses, prospective, cross-sectional between groups study design was completed to address the research questions. Two discourse genres differing primarily along the continuum of memory demands and the degree of visual support provided by the stimulus (i.e., picture sequence description and story retelling) were collected to obtain a large, representative sample of spoken language from participants with PD. An identical set of multi-level discourse measures (i.e., microlinguistic and macrostructural analyses) were used to analyse discourse samples from both genre types for the purpose of developing a profile that would discriminate the discourse performances of participants with PD from that of controls.

Statistical methods were used to optimize the classification of discourse samples by developing a discriminant function, using a minimal set of primal discourse variables that characterized discourse performance differences in PD. A comprehensive battery of standardized assessments was administered to understand how group differences in cognition, language, and speech performance influenced spontaneous language production. Participants with PD were carefully described using disease-relevant measures, detailed self-reports of disease history, and medication profiles. These data, in part, were used to elucidate the nature of any relationships between disease variables and discourse performance. The sections that follow describe the participants and the specific method used to answer the research questions.

3.1 Ethics Approval, Participant Recruitment and Enrolment

Ethics approval. The study received approval from the Health Sciences Research Ethics Board at Western University (Appendix A) and the Lawson Health Research Institutes (Appendix B). All participants provided informed written consent.

Sample size determination. G*Power v3 (Faul, Erdfelder, Lang, & Buchner, 2007) was used to calculate an estimated sample size using the ‘best available’ effect sizes from the literature. No single published study was sufficient for determining effect size estimates because of the unique nature of the protocol. Toward accomplishing the

primary objective of the study, the focus of the power calculation were those analyses used to identify the discourse variables characterising spoken discourse impairment in PD. The power calculation was based on results from the following studies:

- Ash et al. (2012a) reported data from a single discourse task (story generation from a wordless picture book) and a battery of standardized cognitive and language measures using a sample size of 21 PD and 16 healthy controls.
- Murray (2000) reported data from 2 stimuli (picture description) and a battery of standardized cognitive and language measures using a sample size of 10 PD and 9 healthy controls matched to the PD group.

With $\alpha = .10$, the current study ($N = 38$; 19/group) is sufficiently powered to detect large magnitude interaction effects ($f^2 = .40$) for the mixed multivariate analysis of variance (MANOVA) analysis protocol with $1 - \beta = .70$. The power for assessing each MANOVA main effect separately is more robust than for the interaction effect and is able to detect medium effects ($f^2 = .25$) with $1 - \beta = .80$. The univariate F tests are sufficiently powered to assess large effects ($f^2 = .40$) with $1 - \beta = .80$. Given the similarity between the two tests, the power calculation for MANOVA was applied to the discriminate function analysis. An alpha level of 0.10 was chosen for two related reasons: 1) to minimize the risk of not detecting potentially important discriminative discourse variables with smaller effect sizes and 2) to optimize the power of the study given the available resources.

Using the data published in the respective studies, estimated effect sizes for the discourse variables in the Ash et al. (2012) and Murray (2000) studies showed that Cohen's d values ranged from .34 to .77 for PD vs. control differences using discourse measures consistent with those reported in the current study. These data suggest that a wide range of effect sizes are observed in studies of discourse in PD from small-medium to large magnitude effects. Given the wide range of effect sizes the greater concern in the study was failing to find a significant group effect for a discourse variable with a smaller effect size. Therefore the primary goal was to minimize Type II error and optimize $1 - \beta$ to a level of .80 for the main effects testing. To this end, there were three options: modify the statistical procedure, increase sample size, or increase the α level (Cohen, 1992). The

need to analyse between group effects for multiple variables in order to reduce a large set of variables to a minimal set of discriminative variables, as a function of conceptually based discourse domains, necessitated the use of multivariate statistical procedures. Using a multivariate approach increased the power of the analysis via the ‘protected F ’ in the subsequent univariate tests mitigating the need to conduct these with an alpha level corrected for multiple comparison bias (Hummel & Sligo, 1971). Consequently, there was little flexibility around the statistical procedures used. Furthermore, sample sizes in studies of spoken discourse are often smaller (≤ 20 /group), especially true in PD, due in part to recruiting challenges and the labour-intensive and resource-demanding nature of cross-genre, multi-level spoken discourse protocols. Even with a resource-optimized sample size, achieving acceptable power for reducing the risk of Type II error could not be reasonably accomplished without increasing the α level to .10. Therefore to optimize power to $1-\beta = .80$ for the main effect analyses, the alpha level was increased to .10 for the F test procedures used to answer RQ 1 to RQ 4. While a significance level of .10 increases the risk of Type I error this was not of significant concern because in the statistical plan each of the significant variables identified during the MANOVA/follow-up analysis of variance (ANOVA) procedures is subjected to a second analysis for discriminative ability in the discriminant function analysis. In other words, including a variable in the discourse profile that was not ‘truly different’ between PD vs. controls during the first level of analyses (MANOVA/follow-up ANOVA) is mitigated by the second level of analyses (discriminant function analysis), where variables not showing suitable discriminative abilities are typically removed from the discriminant function by the researcher to optimize the predictive accuracy of the function. Determining all potentially discriminative discourse variables using the MANOVA/follow-up ANOVAs within the inherent limitations of the resource demands of this type of research was key for developing an optimal discriminant function. Therefore the resulting $\beta:\alpha$ ratio (Cohen, 1992) of 2:1 (.20 to .10) was deemed acceptable for balancing Type II vs. Type I error risks.

Inclusion/exclusion criteria. The criteria for determining participant eligibility for the study are presented in Table 3.

Table 3

Study Inclusion and Exclusion Criteria

Inclusion	Exclusion
<p><i>All Participants</i></p> <p>English as primary language for speaking and listening in home and work environments.</p> <p>Completed Grade 10 or higher</p> <p>Age 50 to 80 years</p>	<p><i>All Participants</i></p> <p>Positive medical history (self-report and for PD medical record review) for neurological:</p> <ul style="list-style-type: none"> • Injury; • Surgical procedure; • Neurological disease (i.e., acquired or degenerative other than PD). <p>Positive medical history (self-report and for PD medical record review) for:</p> <ul style="list-style-type: none"> • Significant untreated clinical depression; • Other major psychiatric illness (i.e., Schizophrenia, bipolar disorder). <p>Currently taking prescribed cholinergic or psychoactive medications.</p>
<p><i>Control Specific</i></p> <p>Mattis Dementia Ratings Scale – 2: Score \geq 133 or \geq 135 depending on age and education level (Jurica, Leitten, & Mattis, 2001).</p>	
<p><i>PD Specific</i></p> <p>Diagnosis of idiopathic PD made by a movement disorders neurologist.</p> <p>Under medical management of a movement disorders neurologist for \geq 3 years duration.</p> <p>Mattis Dementia Ratings Scale – 2: Score \geq 123/144 (Llebaria et al., 2008).</p> <p>UPDRS-I (item 3) score $<$ 2 (Depression screening item)</p>	

Recruitment and enrolment. PD participants were recruited via two sources: (a) movement disorders specialty clinics located in London, Ontario, Canada and (b) the clinic database of the H.A. Leeper Speech and Hearing Clinic at Western University which provides speech and voice therapy services to client with PD. PD participants were identified as potential participants if they were under the care of a movement disorders neurologist for a duration \geq 3 years, were eligible for the study based on the

inclusion/exclusion criteria when screened by their health care provider, and were willing to be contacted about the study.

Twenty-six individuals with idiopathic PD were contacted about the study of which: (a) two declined the invitation because of the time commitment required; (b) three declined the invitation for non-specific reasons; and (c) one was excluded because she spoke both English and Czech on a daily basis. Twenty individuals with PD met the enrolment criteria and accepted the invitation to participate in the study. Of the twenty participants with PD enrolled, one participant was withdrawn prior to data collection secondary to developing medical complications that affected her cognition. Nineteen participants with idiopathic PD completed the study.

Twenty-one healthy adults meeting the inclusion/exclusion criteria were invited to participate in the study as control participants. Controls were recruited from three sources: (a) family members of individuals with PD, (b) friends of individuals with PD, and (c) community respondents to an invitation to participate in the study. Of those, one person was not enrolled because of a self-reported history of depression discovered during the enrolment process. Twenty healthy older adults without PD met the enrolment criteria and accepted the invitation to participate in the study. One control participant was withdrawn after the start of the study because of extended travel plans that conflicted with the study schedule. Nineteen control participants completed the protocol. See Table 4 (Section 3.2) for a summary of their demographic information.

3.2 Description of Participants

A total of 19 participants with PD and 19 controls completed the study. The participants are described in Table 4.

Table 4.

Demographic Data for Participants by Group

Variable	<i>M (SD), 95% Confidence Interval [Lower, Upper]</i>		Statistical Results
	<i>PD</i>	<i>Control</i>	
Age (yrs.)	70.74 (7.92), 95% CI [66.92, 74.55]	68.16 (7.14), 95% CI [64.71, 71.60]	$t(36) = 1.054, p = .299$
Education (yrs.)	15.47 (2.91), 95% CI [14.07, 16.88]	15.21 (3.36), 95% CI [13.59, 16.83]	$t(36) = .258, p = .798$
MDRS-2 _{Raw} (/144)	140.11 (1.91), 95% CI [139.18, 141.03]	141.95 (2.07), 95% CI [140.95, 142.94]	
MDRS-2 _{MOANS} (/18)	14.25 (1.32), 95% CI [13.58, 14.84]	15.42 (2.19), 95% CI [14.36, 16.48]	$t(29.47) = -2.063, p = .048,$ $d = .67$
<i>PD Only Variables</i>		<i>M (SD), 95% Confidence Interval [x, y]</i>	
UPDRS-III (/108)		30.26 (12.13), 95% CI [24.42, 36.11]	
Hoehn & Yahr (/5)		2.45 (.797), 95% CI [2.06, 2.83]	
Duration of PD symptoms (yrs.)		9.34 (3.63), 95% CI [7.59, 11.09]	
Levodopa equivalent dose mg (LED)		984.37 (539.95), 95% CI [724.12, 1244.62]	

Note. $*\alpha = .05$ for Age, Education, and MDRS_{MOANS} t-tests. Equal variances assumed for Age and Education as Levene's tests were not significant ($p > .05$) for Age or Education.

** Levene's test for the MDRS_{MOANS} was significant using a significance level of .05 ($F(1, 36) = 16.95, p = < .000$) so the t values for the adjusted df were reported.

There is an asymmetric distribution of sex across the groups with men more heavily represented in the PD group (Men = 15) than in the control group (Men = 7). The sex distribution of the PD group fits a similar pattern to the typical demographic profiles wherein the incidence of PD is reported as 1.5 times higher in men vs. women and

prevalence rates are reported as almost two times higher in men vs. women (Pavon, Whitson, & Okun, 2010; Wooten, Currie, Bovbjerg, Lee, & Patrie, 2004). Notwithstanding the similar pattern of expected sex distribution, the ratio of men to women in the PD group is slightly skewed toward men. However, gender effects are negligible for spoken discourse using highly structured discourse tasks such as those reported in the current study (Mackenzie, 2000). Moreover, gender effects are not significant for those standardized cognitive and language measures, identical to or similar to those administered in the protocol, for either healthy adults (Parsons, Rizzo, Zaag, McGee & Buckwalter, 2005; Strauss, Sherman & Spreen, 2006; Zec, Burkett, Markwell & Larsen, 2007) or for individuals with PD (Elgh et al., 2009). As such the asymmetry of gender distribution between groups is unlikely to have affected data collection, data analysis, or interpretation of the results. With a significance level of .01, a non-significant Shapiro-Wilk's test, suggested both PD and control data sets were normally distributed for both age (PD, $p = .032$; control, $p = .634$) and education (PD, $p = .110$; control, $p = .201$). There were no significant differences in age or years of formal education between groups based on separate individual groups t-tests with a significance level of .05 (see Table 4).

However, there were significant differences between groups for global cognition using the Mattis Dementia Rating Scale – 2 (MDRS-2). The MDRS-2 control group data violated assumptions of normality ($p < .001$) using the Shapiro-Wilk's test with a significance level of .01; however, the PD group data were normally distributed (Shapiro-Wilk's $p = .556$). Given the robustness of t-tests to violations of normality when n sizes are equal, this violation was not of significant concern (Khan & Rayner, 2003). There was a significant PD vs. control difference for the MDRS-2 MOANS (Mayo's Older Americans Normative Studies) age adjusted scaled scores (Lucas et al., 1998) (Table 4). A Cohen's d value of 0.67 for the effect of group on global cognition was interpreted as a medium effect size (Cohen, 1988). Importantly the differences in mean scores between the two groups was < 2 points. A score change of 2 points or greater is suggested to indicate a meaningful clinical difference on the MDRS-2 (Jurica et al., 2001). So while the MDRS-2 scores did significantly differ between groups, this difference was not one

of significant clinical meaningfulness. Individuals with PD may perform more poorly than healthy adults on dementia screening tools for reasons unrelated to dementia or mild cognitive impairment, including motor and visual perceptual impairments, that affect time constrained and construction tasks (Kalbe et al., 2008). Consequently, the significant t-test for the MDRS-2 was followed with an examination of the PD participants' performance using disease specific normative data for the MDRS (Llebaria et al., 2008). Each of the individual scores for the PD participants was above published MDRS cut-off scores for dementia (i.e., $\geq 123/144$). Importantly, using the normative data from Llebaria et al. (2008) to convert the MDRS-2 raw scores into z-scores all of the standardized scores fell above the mean ($M_{\text{MDRS z-score}} = 1.02$, $SD = .295$, Range = .37 to 1.62, 95% CI [.87, 1.16]) suggesting, with confidence, that none of the PD participants met the criteria for dementia. A wide range of UPDRS-III scores, Hoehn and Yahr scores, duration of PD symptoms, and levodopa equivalent dose (LED) characterized PD participants, reflecting the variability observed in the PD population at large (van Rooden et al., 2011). All of the PD and control participants self-reported right-hand dominance save one control participant who reported left hand dominance. All participants self-reported English as their primary language of use in daily activities. Only one participant described himself as bilingual being raised in a bilingual home (French and English) but used English exclusively as his primary language for more than 40 years. No participants reported a diagnosis of untreated depression or major psychiatric illness. A review of the medical records for the participants with PD confirmed the absence of untreated depression. For the 'Depressed Mood' item of the UPDRS-I, scores ranged from 0 to 1 (max possible score = 4) with a median score of 1 and a mode of 0.

3.3 Audio Recording Protocol

For 37 of 38 participants, spoken language tasks and speech tasks were recorded using an AKG C520 head-worn condenser cardioid microphone (i.e., manufactured by AKG Acoustics GmbH Vienna, Austria). The headset was positioned comfortably allowing hearing aids and prescription eyewear to remain in place. The microphone was positioned such that the signal to source recording distance was 2 to 4 cm, depending on head size/jaw length, allowing for limitations in the microphone arm length. The

recordings were made with the microphone connected to a Macbook Pro laptop (i.e., OS X version 10.9 Intel Core i7 8 GB 1600 MHz DDR3 manufactured by Apple Inc. Cupertino, California USA) via an USBPre2 audio interface with 48 V Phantom Power (i.e., manufactured by Sound Devices, LLC Reedsburg, Wisconsin USA). Audio data were recorded using Audacity® 2.0.5 (<http://audacity.sourceforge.net>), a free open source software for recording and editing. For recording, the project rate was set to 44100 Hz, 16-bit, mono input, with the USBPre2 set as the input and output sources. Files were saved as .wav files. Each task in the protocol was saved as a separate file. The microphone gain was customized for each participant using a test protocol. For the test protocol a test speech sample was recorded while the participant counted aloud from 1 to 10. While the participant counted aloud, the signal output in Audacity® was continuously evaluated for evidence of peak clipping. Simultaneously, the gain monitor on the USBPre2 audio interface was monitored and adjustments were made, as needed, to ensure the signal recorded was of sufficient quality per the manufacturer's instructions. For 1 of 38 participants, a non-manufacturing related microphone failure occurred during the recording equipment set-up necessitating use of a backup recording system. For this participant, the data were collected using a VIXIA HF M500 HD Camcorder (i.e., manufactured by Canon, Inc., Tokyo, Japan) fitted with a Røde VideoMic Pro shotgun condenser cardioid external microphone (i.e., manufactured by Røde Microphones Sydney, Australia). The microphone was connected to the video recorder using an extension cable and was affixed to a JOBY GorillaPod mini tripod (i.e., manufactured by JOBY, Petaluma, California USA) positioned on a table such that the signal to source recording distance was 30 cm. The Camcorder was affixed to a separate standard-sized Velbon aluminum tripod (i.e., manufactured by Velbon, Maidenhead, Berkshire, UK) and was turned away from the participant's face such that audio data only were recorded directly from the participant. For this one recording, the files were saved as .MTS files.

3.4 Data Collection Environment

Participants completed the study protocol either in their home or in the Aging and Communication Disorders Laboratory located in Elborn College at Western University. The bulk of data collection occurred during the winter months in southwest Ontario,

Canada. Consequently, 17 participants with PD and 15 control participants elected to complete the study protocol in their homes to avoid exposure to adverse weather conditions. The six remaining participants elected to complete the protocol in the laboratory because it was more convenient to do so for a variety of reasons including: (a) a terminally ill family member under 24-hour care in the home, (b) un-removable distractions such as a one spouse caring for a grandchild or home renovations, and (c) convenience around work or other scheduled activities. For the in-home data collection, participants were seated in a comfortable chair typically at the dining room/kitchen table. Conditions for completing the study protocol were optimized, specifically: (a) televisions and radios were turned off, (b) lighting was optimized for completing tasks, (c) only the researcher (AR) and the participant were in the testing room/area, and (d) other participant-specific environmental distractors were removed or minimized as indicated (e.g., pets). In the laboratory environment, similar conditions occurred with the researcher (AR) and the participant positioned at a table, lighting optimized, and distractions removed or minimized. The recording set up was the same for both environments. The specifications of the recording equipment, particularly the microphone, minimized concerns with environmental effects on the audio recordings.

3.5 General Method

All data were collected, scored (i.e., with the exception of the measure of speech intelligibility), and analyzed by a single researcher (AR). She is a registered speech-language pathologist experienced in the administration, analysis, and interpretation of the tasks/assessment measures included in the protocol.

Data from participants with PD were collected in their optimal ‘on’ state relative to their PD medications. PD participants selected their ‘best’ appointment time of the day relative to their typical cognitive status, fatigue levels, motor performance and medication schedules (i.e., typically mid-morning) at the time of research appointment booking. Medication schedules were not altered for the study. Medication schedules were variable across participants. However, when possible, research visits were scheduled in the window between 30 minutes post PD medication intake and 30 minutes

prior to next dosing. A break was provided such that participants could take their medications during the data collection session when required by their typical medication schedule ($n = 2$).

The study protocol was completed in a single session with an average duration of 2.75 hours (Range = 2.5 to 3.5 hours). Participants were provided rest breaks every 60 minutes or as requested. However, most participants did not find the rest breaks necessary. The protocol was tolerated well. All participants ($N = 38$) completed the full protocol without voiced concerns of fatigue.

An overview of the study protocol including the presentation order of tasks is presented in Table 5. A detailed description of the study tasks including administration and scoring procedures follows in Section 3.6 to 3.12. The order of administration of all study tasks remained fixed to maintain the consistency of any order effects across participants because of the potential for such (i.e., interference or facilitation) on the standardized measures of cognition, language, and speech intelligibility. Items more sensitive to fatigue (i.e., speech intelligibility, cognitive measures of attention and interference) were administered earlier in the data collection session to optimize performance. Study tasks more sensitive to familiarity with the researcher (i.e., discourse tasks) were performed later in the data collection to optimize performance via established rapport with the researcher and increased familiarity with the study protocol. A total of eight discourse tasks (i.e., 2 stimuli each from 4 genres) were administered, four of which are relevant to the proposed research questions and are presented herein. To avoid order effects in the discourse data, the discourse stimuli were presented in a randomized order (i.e., across all 8 stimuli) using a randomization schedule developed prior to initiating the study (using Random.org <http://www.random.org/>).

Table 5

Overview of Study Tasks and Order of Administration

Order	Task	Group(s) Completing
1	Cognitive screening - Mattis Dementia Rating Scale – 2 (Pre-screening eligibility)	Participants without a current (i.e., within 6 months), verifiable score in their clinic/research records

2	Consent	All
3	Hearing screening	Participants not currently wearing amplification
4	Vision screening	All
5	Demographic interview	All
6	Grandfather Passage Reading	All
7	Sentence Intelligibility Test	All
8	Stroop Color-Word Test	All
9	Trail Making Test (A & B)	All
10	Boston Naming Test	All
11	Verb Picture Naming Subtest: Test of Adolescent and Adult Word Finding	All
12	Pyramids and Palm Trees	All
13	Verbal fluency tasks (Semantic, Phonemic, Action)	All
14	Discourse tasks: Set 1 (4 tasks)	All
15	Sentence Production Priming Test from the Northwestern Assessment of Verbs and Sentences	All
16	Discourse tasks: Set 2 (4 tasks)	All
17	Verbal learning subtest (Arizona Battery of Communication Disorders of Dementia)	All
18	Disease related measures (i.e., UPDRS-III and Hoehn & Yahr)	PD only
19	Medication interview & verification	PD only

All of the spoken language tasks, including the discourse tasks and the standardized cognitive, language, and speech intelligibility measures were audio recorded using the methods described in Section 3.3. The audio recordings served as the primary data source for the discourse data and for the speech intelligibility data. The audio files were used off-line to verify responses and accuracy of on-line scoring for the standardized measures of cognition and language.

3.6 Hearing Screening

Participants who, at the time of the study, were not currently wearing a hearing amplification device (e.g., hearing aid(s)) completed a hearing screening protocol (34 completed; 4 participants had existing amplification) to ensure sufficient hearing acuity for completing study tasks. Pure tone hearing screenings were conducted in accordance with the American Speech-Language-Hearing Association *Guidelines for Audiologic Screening* for adults (ASHA, 1997). A single, calibrated, GSI-18 Screening Audiometer (i.e., manufactured by Grason-Stadler Incorporated, Eden Prairie, MN, USA) with TDH-39 headphones was used for screening. Six participants failed the hearing screening and were referred for further audiologic testing. Participants who failed the hearing screening were fitted with a Bellman Audio Maxi Personal Amplifier (i.e., manufactured by Bellman & Symfon, Gothenburgh, Sweden). When required, personal amplifiers were worn only during the administration of task instructions, and removed during speech production and spoken language tasks. A test of comfortable loudness was performed using the personal amplifier. The researcher, seated in the same position in which data were collected, counted aloud from 1 to 20, with mouth visually occluded, using normal conversational speech volume. Participants self-adjusted the volume of the personal amplifier, as the researcher counted aloud, until the loudness level was comfortable.

3.7 Vision Screening

All participants were asked to complete a vision screening (38 completed) to ensure sufficient visual acuity for completing the study tasks. Participants were encouraged to wear their normal prescription eyewear, if typically used for reading, both during vision screening tasks and during data collection. The vision-screening subtest from the Arizona Battery of Communication Disorders of Dementia (Bayles & Tomoeda, 1993) was administered. The vision screening consisted of three tasks: (a) reading aloud 2 sentences printed in 14 point black sans serif font in portrait page orientation to assess for visual acuity/literacy, (b) completing a visual field letter cancellation task with five target letters per quadrant to assess for visual/spatial neglect, and (c) naming two black and

white line drawings of high frequency objects presented one at a time to assess for visual agnosia. All participants (N = 38) passed the vision screening.

3.8 Collection Demographic Information

Participants were asked to provide basic demographic information as a part of the study (38 complied). The demographic information collected included: (a) age, (b) dominant hand of use for writing and for activities of daily living (i.e., handedness), (c) language(s) for which the participant was competent both in understanding and in speaking which they used at least weekly to communicate either at home or at work (i.e., language), and (d) the number of formal years of education completed (i.e., education). All measures were collected via self-report. See Table 4 (Section 3.2) for a summary of these data.

3.9 Collection of Disease Related Measures

Participants with PD were asked to complete two standardized measures specific to PD: a) one for motor severity the UPDRS-III (Fahn et al., 1987) and b) one for disease staging the modified Hoehn and Yahr (Goetz et al., 2004) (19 completed). The researcher (AR), who is trained and certified, administered both the UPDRS-III and Hoehn and Yahr scales to all participants to optimize inter-subject reliability. The UPDRS-III and Hoehn and Yahr scales were administered on the same day (i.e., within the same session) as the cognition, language, speech, and discourse measures were collected. Additionally, PD participants were asked to provide information relative to their disease onset and medications. Participants self-reported the number of years they had experienced PD symptoms (i.e., duration) (19 complied). PD participants were also asked to provide a list of medications including dosing information. Medications were subsequently verified by the researcher via a review of the prescription labels on their medication containers (19 complied). PD medications were converted into and subsequently recorded as levodopa equivalent doses (LED) using the conversion formulae provided by Tomlinson, et al. to facilitate comparison of medications across participants (Tomlinson et al., 2010). See Table 4 (Section 3.2) for a summary of these data.

3.10 Measures of Cognition and Language

The battery of cognitive and language measures developed for the study protocol was informed by: (a) the existing research literature examining cognition and language performances in PD, (b) the current body of spoken monologic discourse literature in PD, (c) the clinical experience of the researcher, and (d) the published protocols for the neuropsychological assessment of cognitive decline in PD (Litvan et al., 2012). The goal was to elucidate a comprehensive profile of cognitive and spoken language abilities that could affect discourse performances. Measures of language comprehension and written language expression were excluded purposefully because the focus of the research was spontaneous spoken language performance. The battery administered in the study protocol and the constructs assessed by each standardized measure are presented in Table 6 followed by a more detailed description of these measures and scoring procedures.

Table 6

Standardized Measures of Cognition and Language Administered in the Study Protocol

Measure	Construct(s)
Boston Naming Test – 2 nd edition (30-item version) (Kaplan, Goodglass, & Weintraub, 2001)	<i>Language.</i> Spoken confrontation naming for nouns manipulated by word frequency
Picture Naming (verb) subtest of the Test of Adolescent and Adult Word Finding (German, 1989)	<i>Language.</i> Spoken confrontation naming for verbs manipulated by word frequency
The Pyramids and Palm Trees Test (3-picture version) (Howard & Patterson, 1992),	<i>Language.</i> Semantic knowledge/memory aspects of word retrieval
Sentence Production Priming Test (Thompson, 2011)	<i>Language.</i> Spoken syntax production for both simple and complex (i.e., embedded clause), canonical and non-canonical word order sentences
Verbal Fluency (Semantic -Animals)	<i>Language.</i> Timed spoken word retrieval from semantic categories (i.e., living creatures)

Verbal Fluency (Phonemic – F, A, S)	<i>Executive Function.</i> Timed spoken word retrieval for words by initial phoneme
Verbal Fluency (Action)	<i>Language & Executive Function.</i> Timed spoken word retrieval from a semantic category mediated by prefrontal cortex (i.e., actions)
Stroop Color-Word Test (Golden & Freshwater, 2002)	<i>Frontal Lobe Function.</i> Inhibition specifically the ability to maintain a goal and suppress an entrained response
Trail Making Test (United States Army, 1944; Reitan & Wolfson, 1985)	<i>Frontal Lobe Function.</i> Attention, speed, mental flexibility
Verbal learning subtest of the Arizona Battery of Communication Disorders of Dementia (Bayles & Tomoeda, 1993)	<i>Episodic Memory.</i> Specifically verbal learning and memory.

A validated 30-item version of the Boston Naming Test – Second Edition (BNT) was used (Graves, Bezeau, Fogarty, & Blair, 2004). The stimuli plates, instructions and scoring procedures were implemented without modification from the BNT. The picture naming (verb) subtest of the Test of Adolescent and Adult Word Finding (TAWF-verb), The Pyramids and Palm Trees Test, The Sentence Production Priming Test from the Northwestern Assessment of Verbs and Sentences (SPPT), and the verbal learning subtest of the Arizona Battery of Communication Disorders of Dementia (ABCD-VL) were administered and scored according to their respective published manuals.

Both the BNT and the TAWF-verb required participants to look at line drawings of either objects (i.e., 30 for the BNT) or of actions (i.e., 25 for the TAWF-verb) and to produce a single word that referred to the object/action presented in the pictures. Both measures assess confrontation naming (BNT for pictured nouns and the TAWF-verb for pictured verbs). For The Pyramids and Palm Trees test, participants looked at 52 panels of black and white line drawings (i.e., three drawings per panel with the target picture on the top row and the two picture choices positioned beneath it). Participants identified which drawing out of two choices ‘best matched’ the third line drawing (i.e., the target).

Correct responses are dependent on preserved semantic memory and the ability to use visual features from the pictured objects to access the semantic system and then use that information to determine which two concepts, represented by pictures, are more closely associated. The ability to access and utilize semantic information is a critical component of contextual word retrieval in spontaneous language. The SPPT is designed to systematically test for productive syntax abilities controlling for lexical retrieval demands. For the SPPT participants were presented with two black and white line drawings of scenes depicting actions between two or three actors (i.e., the number of actors was dependent on the syntax structure elicited). The researcher provided a model sentence using a specific syntax structure (e.g., object-relative) for one of the pictures while identifying the actor(s) in the scene and the action being performed. Subsequently, participants generated a novel sentence describing the events of the second picture (same actors in different roles; same events) using the identical syntax structure modeled by the researcher. The SPPT contains 30 sentences in total, 5 each from 6 different syntax structures including sentences with embedded clauses, Wh- questions, and sentences with canonical and non-canonical word ordering. The ABCD-VL uses a word list-learning paradigm (i.e., 16 words) to assess both uncued and cued delayed verbal recall. While both the correct number of cued and uncued recall items were recorded, only the uncued score was used in the statistical analysis because it held the greatest relevance to the current protocol.

The semantic verbal fluency (i.e., Semantic) and phonemic verbal fluency (i.e., Phonemic) tasks required participants to produce as many spoken exemplars as possible from their respective task-specific categories within 60 seconds duration (i.e., animals – Semantic or words starting with a specified letter – Phonemic). While semantic fluency reflects semantic processes, phonemic fluency is often associated with executive functions (Henry & Crawford, 2004). Both tasks were administered and scored according to published procedures (Spreen & Strauss, 1998). While action fluency performance may reflect a specific disruption in verb access in PD and was included for that reason, it is also used to assess processes of executive function mediated by the frontal lobes (Piatt, Fields, Paolo, & Tröster, 1999b). For the action verbal fluency (i.e.,

Action) task participants generated as many action verbs (i.e., “things people do”) as they could within 60 seconds duration. Action fluency was administered and scored according to published procedures (Woods et al., 2005).

The Stroop Color and Word Test – Adult (i.e., Stroop) 45-second version, a measure of executive function, was administered using the stimuli, instructions and scoring rules published in the test manual for clinical and experimental use authored by Golden and Freshwater (2002). The Stroop is a speeded reading task that uses three sequentially presented stimulus cards each with a 5 x 20 matrix of printed words or letter strings: (1) one card with color words printed in black ink (i.e., presented first to participants), (2) a second card with a series of letter strings ‘XXXX’ printed in either red, blue, or green fonts, and (3) a third card with color words printed in ink colors that conflict with the printed word (e.g., the word *blue* printed in red ink **BLUE**) which is presented last to participants. For the first card, participants read the printed words aloud as quickly, but as accurately, as possible proceeding from the top to the bottom of each column for a 45-second duration. For the second card, participants state the color ink in which the letter string is printed. The task is identical for the third card except participants must suppress reading the word in order to state the color of the ink in which the word is printed. This ‘interference’ effect is unique to color naming and is an established measure for assessing “interference in verbal processing” (p. 2) (Golden & Freshwater, 2002). The 45-second version has been shown to be a more stable measure and less fatiguing for individuals with neurological injury and thus was used for the current protocol (Golden & Freshwater, 2002). For the purposes of this study, the published normative data from the test manual were used to calculate the interference score, which was the score reported and the score used in statistical analyses.

The Trail Making Test (Trails) is a timed pen and paper task consisting of two separate and sequentially administered parts (Trails A and Trails B) (United States Army, 1944; Reitan & Wolfson, 1985). It is designed to test attention-switching, speed of cognitive processing, and executive functions. In each part, using a continuously drawn line participants connected (i.e., as quickly but as accurately as they could) twenty-five labeled circles in consecutive ascending order using either numbers only (i.e., Trails A)

or alternating between numbers and letters (i.e., Trails B). Trails A and Trails B were administered using the instructions published in (Spreen & Strauss, 1998). For between group comparisons, the score reported and used in the statistical analysis was the ratio of the time to complete Trails B and the time to complete Trails A (Trails B/A). The scoring method is reported to reflect accurately executive process and attention switching, controlling for motor speed and visual scanning speed differences across participants; an important consideration given the inherent motor deficits in PD (Arbuthnott & Frank, 2000).

3.11 Measures of Speech Intelligibility and Rate

Speech and voice changes are ubiquitous in Parkinson disease (Ramig et al., 2004). Two measures of motor speech performance were administered, one for intelligibility and one for rate, in order to assess the impact of group differences in speech and voice performance relative to the integrity of and interpretation of the discourse data. The methods for collecting these measures are detailed herein.

Measure of speech intelligibility. Each participant completed the 11-sentence version of the Sentence Intelligibility Test (SIT), a subtest of the Assessment of Intelligibility of Dysarthric Speech (Yorkston et al., 1984). The SIT is a transcription measure of speech intelligibility in which naïve scorers listen to recorded samples of clients', and in this case participants' speech, and transcribe orthographically the words they hear in the recordings. As a part of the test protocol, naïve listeners hear a recorded speech sample twice, once without stopping the recording to acclimatize to the speaker's voice/speech pattern (i.e., no transcribing) and once to transcribe. During the transcription stage, listeners are able to stop the recording as often as necessary to transcribe the sample, but are not allowed to replay the recording to verify or clarify a transcription. Speech intelligibility for the SIT is calculated by dividing the number of correctly transcribed words by the total number of words in the sample. Thirty-eight different, randomized, 11-sentence stimuli sets consisting of sentences 5 to 15 words in length (i.e., one sentence of each length) were generated using the Sentence Intelligibility Test for Windows (Yorkston, Beukelman, & Hakel, 1996) and were randomly assigned

to participants. The sentence sets were printed on a single sheet of 8 x 11 inch paper in landscape page orientation in 16 point Arial black font. Participants positioned the paper comfortably on the table in front of them such that they could easily see all of the sentences. Once the paper with the printed sentence stimuli was positioned the instructions were provided and audio recording initiated. Using the verbatim instructions provided for the SIT in the Assessment of Intelligibility of Dysarthric Speech manual, participants were instructed to read each sentence aloud, using their normal ‘voice volume’ and ‘speaking rate’, one at a time in sequential order until they completed reading all of the sentences (Yorkston et al., 1984).

SIT scoring. The SIT was transcribed and scored according to the published instructions in the manual (Yorkston et al., 1984). SIT samples ($N = 38$) were scored separately by two scorers (i.e., volunteer graduate students in audiology and speech-language pathology enrolled in the School of Communication Sciences and Disorders). Five separate scoring sessions (i.e., 90 to 120 minutes duration) were conducted in the Health and Rehabilitation Sciences Computer Lab at Elborn College, Western University. Each scorer participated in a 30-minute training session (i.e., conducted by the researcher) prior to completing independent scoring of samples. The training session included a description of the SIT measure, description of the data collection methods, and a review of the instructions with transcription/scoring examples.

Each scorer was provided with a written copy of the scoring instructions, abstracted from the test manual (Yorkston et al., 1984), for their reference during the transcription and scoring processes along with scoring sheets for recording the transcriptions. The audio recordings of the SIT samples were assigned randomly to each scorer. All audio recordings were fully blinded (i.e., scorers did not know if the samples were PD or control). Each scorer sat at a separate computer terminal. Scorers listened to SIT audio recordings with Windows Media Player (manufactured by Microsoft, Redmond Washington, USA) using JVC HA-S160-V headphones (manufactured by JVC Kenwood Corporation, Yokohama, Japan). After transcribing their randomly assigned samples, scorers exchanged score sheets with a second scorer who was naïve to the source SIT recordings. The second scorers used the ‘answer key’ provided (i.e., the stimulus sheets

from which participants read the sentences) to calculate the number of correctly transcribed words. SIT scores were reported as percent of correctly transcribed words (i.e., intelligibility). This procedure was followed twice, with two different raters, for each SIT recording generating two SIT scores for each sample. The final score entered into statistical analysis was the average of these two scores.

Agreement study. Twenty percent of the SIT samples were selected randomly for an intra-rater agreement study in which scorers re-transcribed a previously heard sample. An equal proportion of PD and control samples were included in the agreement study. The first and second transcriptions were separated by at least 72 hours. For the inter-rater agreement study, a second blinded scorer transcribed a separate randomly selected sample of twenty percent of audio recordings. For both the intra-rater and inter-rater agreement studies rater agreement was assessed using the intraclass correlation coefficient (ICC). ICCs are widely used in the rehabilitation literature for reliability/agreement studies where interval level data are collected (Rankin & Stokes, 1998). Intra-rater agreement for the SIT data was interpreted as excellent (Cicchetti, 1994) with $ICC(1, 1) = .90$, 95% CI [.710, .963]. Inter-rater agreement was less robust with $ICC(3, 1) = .53$, 95% CI [- .240, .804] interpreted as fair agreement (Cicchetti, 1994). The inter-rater reliability for the SIT was lower than previously reported values from other studies in which experienced speech-language pathologists (SLPs) vs. novice student clinicians completed the SIT transcription scoring. Constantinescu et al. reported an ICC value of .94 for inter-rater reliability on the 22-sentence version of the SIT for three experienced SLP raters using speech samples produced by PD participants (Constantinescu et al., 2010). Whether the inter-rater agreement differences result from rater experience levels of the raters (i.e., novice vs. experienced), the number of items administered (i.e., 11 vs. 22), and/or number of raters (i.e., 3 vs. 2) is unclear from the present data and the existing literature. However, the less robust inter-rater agreement data reinforced the importance of entering an average score from two raters into the analysis as a way of partially, although not fully mitigating, issues with systematic measurement error.

Measure of Speech Rate. It was important to collect measures of baseline speech rates for PD and controls because multiple discourse measures were based in a rate unit (i.e., words per minute). PD affects speech rate resulting in reduced speech rate for some individuals and marked increases in speech rate (i.e., tachyphemia) for others (Ramig et al., 2004). A speech sample was collected from each participant using the audio recording protocol outlined in Section 3.3. Participants read aloud the Grandfather Passage (Darley, Aronson, & Brown, 1975), a 130- word passage used ubiquitously for assessing reading and speech intelligibility. Speech rate was calculated based on the length of the speech sample (i.e., seconds) starting from the onset of the first articulated speech sound of the first word of the passage until the termination of the final articulated speech sound of the final word of the passage using Audacity® 2.0.5 (<http://audacity.sourceforge.net>). The duration of the speech sample was divided by the total number of words (130) and then multiplied by 60 to convert to a speaking rate value in words per minute.

3.12 Elicitation of Discourse Samples

The discourse methodology and data presented are components of a larger ongoing study of discourse in PD conducted by the researcher. For clarity, the discourse task protocol for the larger ongoing study is outlined in Appendix C. However, only two of these tasks are relevant to the current research questions and are discussed herein.

Participants in the current study were asked to produce a total of four spoken discourse samples (i.e., 2 picture sequence descriptions and 2 story retellings). It was necessary to collect a large discourse sample (i.e., 300 to 400 words) that captured the inherent variability across discourse genres while preserving the ability to apply an identical battery of discourse measures to each task. This was undertaken with the ultimate goal of creating a composite variable that could effectively characterize the unique discourse profile of individuals with PD. Picture sequence description (i.e., Picture) and story retelling tasks (i.e., Retell) were used because they are distinguishable by the memory demand and visual support differences between the two tasks. The difference creates a continuum of discourse tasks to include in a cross-genre protocol. Both tasks encouraged

the use of story schemas, required a high degree of lexical specificity, and presented similar syntax demands in that they contain similar numbers of actors and events. However, they differed critically in the fact that the Picture task was a ‘maximum support task’ meaning that the participant was able to use the visual stimulus to support retrieving words and generating spontaneous spoken language with minimal demands on memory. Moreover, these two highly structured elicitation tasks (i.e., Picture and Retell) minimize issues with off-topic or tangential language, a concern relative to age-related differences in discourse performance (Marini et al., 2005). In contrast, the Retelling task provided no visual support and presented maximal memory demands for producing spontaneous language specific to the task. These two genres are well suited for conducting a multi-level, cross-genre discourse study because identical discourse analyses could be applied to both tasks (Marini et al., 2005).

Picture sequence description. For the picture sequence descriptions, two stimuli (i.e., ‘Argument’ and ‘Directions’) published by Nicholas and Brookshire were used (Nicholas & Brookshire, 1993). The stimuli are presented in Figures 1 and 2, respectively.

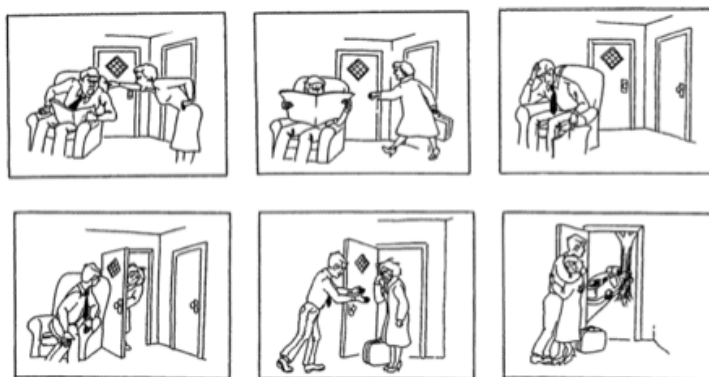


Figure 1. Picture sequence description 1: ‘Argument’. © 1992 Linda Nicholas and Robert Brookshire. Used and reproduced with the permission of Linda Nicholas the surviving author and copyright holder.

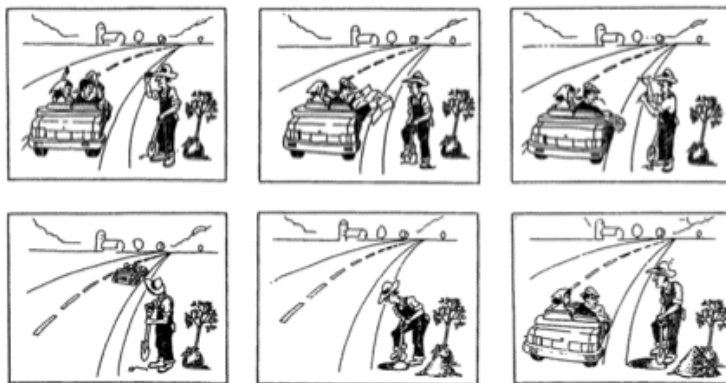


Figure 2. Picture sequence description 2: ‘Directions’. © 1992 Linda Nicholas and Robert Brookshire. Used and reproduced with the permission of Linda Nicholas the surviving author and copyright holder.

Each picture sequence (controlled for equivalent image sizes) was presented on a single 8 x 11.5 inch piece of paper printed in landscape orientation. Each of the picture sequences, depicting a series of events occurring between a set of actors (i.e., 2 actors in the ‘Argument’ picture sequence and 3 actors in the ‘Directions’ picture sequence), contains six black and white, chronologically-ordered, line drawings of scenes. These previously validated stimuli generate reliable and stable discourse data sets across multiple administrations (Brookshire & Nicholas, 1994b). They are reported extensively in the discourse literature for use among both healthy controls and persons with acquired communication disorders (Brookshire & Nicholas, 1994b; Brookshire & Nicholas, 1994a; Capilouto et al., 2005; McNeil et al., 2007; Nicholas & Brookshire, 1993; Nicholas & Brookshire, 1995).

The Picture stimulus was positioned such that the participant could comfortably visualize all of the images in the sequences. Participants adjusted the position of the stimulus as needed prior to starting the discourse production component of the task. The stimulus remained in place until the participant finished producing the discourse sample. The participants had 60 to 75 seconds to review the picture stimulus prior to producing the discourse sample. Participants heard the following instructions: “I am going to ask you to tell me a story about these pictures. First I want you to take some time to look at this

series of pictures and to familiarize yourself with them.” Once the review time had elapsed, the following instructions were provided and the audio recording initiated: “I want you to tell me a story about the people and events you see in these pictures.” No feedback was provided. However, the Researcher provided minimal back-channel prompts for continuation such as “uh huh” and “hmm”. If samples were < 45 seconds duration, participants were provided a single prompt: “Can you tell me more?” These instructions are consistent with those published (Brookshire & Nicholas, 1994b).

Story retelling. For the story re-telling tasks, two stimuli from the Discourse Comprehension Test (DCT) (Brookshire & Nicholas, 1997) were randomly selected, one from Story Set A “The Glass of Water” and one from Story Set B “Out of Gas”. Transcripts of the two story retelling stimuli are presented in Figures 3 and 4, respectively. Written permission was obtained from Nicholas, the surviving author of the DCT and Porch, the publisher of the DCT to use and to reproduce the Retell stimuli.

Joe and Betty Adams were sitting in the living room of their small house. Joe had just finished shovelling the walk. Now he was sitting with his shoes off and his feet up on the footstool watching TV. Betty was busy knitting a sweater for their grandson who was going to graduate from college next month. Joe looked over at his wife and asked her if she would bring him a sandwich. Betty put down her knitting and went into the kitchen. Soon she came back with a cheese sandwich on a plate. When the sandwich was gone, Joe asked his wife to bring him some of the cookies that she had baked that afternoon. Betty went into the kitchen and came back with the cookies and put them on the table beside her husband. Then she went back to her knitting. In a few minutes, Joe said that he was thirsty and asked Betty to bring him a glass of water. Betty looked over at him and said, “I really think that you should get things for yourself once in a while.” Joe slowly got up and went into the kitchen. Soon he reappeared in the doorway and said, “Dear, where do we keep the water?”

Figure 3. Transcript of story retelling 1: ‘The Glass of Water’ (Brookshire & Nicholas, 1997). © 2008 BRK Publishers. Used and reproduced with the permission of Linda Nicholas and Bruce Porch, the surviving author, publisher, and copyright holders.

Jim Hanson was a traveling salesman who sold paint to hardware stores throughout the state of Iowa. Shortly after breakfast one day, he was driving through the countryside when his car gave a sputter and died. He got out and looked under the hood, but he couldn't see anything wrong. He sat down under a tree beside the road for about half an hour, but nobody came by. Finally he decided to try again to start his car. He was muttering to himself about cars always breaking down as soon as the warranty expired. Suddenly he noticed that the needle of the gas gauge was resting on the red "E." Swearing to himself, he opened the trunk and got out an empty gas can. Then he started walking to a gas station. After about a mile, he saw an old man standing beside the road. He stopped and asked the man how far it was to the nearest gas station. The man thought for a minute. Finally he said, "Oh, I'd say a couple of miles, as the crow flies." The salesman wiped his sweaty forehead, and asked, "Well, how far is it if the crow is walking and carrying a gas can?"

Figure 4. Transcript of story retelling 2: 'Out of Gas' (Brookshire & Nicholas, 1997).

© 2008 BRK Publishers. Used and reproduced with the permission of Linda Nicholas and Bruce Porch, the surviving authors, publisher, and copyright holders.

The stories from the DCT are controlled for a number of variables relevant to the current protocol including: number of words, number of sentences, mean length of sentence, number of subordinate clauses, number of utterances, ratio of clauses to T-units, listening difficulty, and number of low frequency words of the DCT manual (Brookshire & Nicholas, 1997, p. 7). Stimuli were presented via digital audio recording using the test materials provided with the DCT. Using Audacity® 2.0.5 (<http://audacity.sourceforge.net>) stimuli were played through the internal speakers of a Macbook Pro laptop (i.e., described earlier in 3.4) positioned at a source to listener distance of approximately 3 feet. The digitally recorded stimuli provided with the DCT were recorded in a male speaking voice using "normal stress and intonation" (Brookshire & Nicholas, 1997, p. 12 of the test manual) with a mean speaking rate of 133 WPM (Range = 128 to 140 WPM). Prior to playing the stimuli for the participant, the researcher performed a test of comfortable loudness. A 20 to 30 second sample from the DCT practice stimuli (P2: The Storm, not used in the study protocol) was played while the output speaker volume was adjusted until the participant verified that a comfortable listening level had been reached. Following the test of comfortable loudness participants heard the following instructions: "You are about to hear a short story. Listen to the story carefully. When the story is completed you will be asked to retell the story." (Doyle et al., 1998, p. 571). After listening to the audio recording of the story the following prompt was provided: "Retell the story in your own words." (Doyle et al., 1998, p. 572).

Participants heard each story once. Participants retold the story immediately after listening to the recording. No feedback was provided. However, the Researcher provided minimal back-channel prompts for continuation such as “uh huh” and “hmm”. If samples were < 45 seconds duration, the researcher provided a single prompt: “Can you tell me more?” The stimuli, instructions and procedures used in the current study are consistent with those published by Doyle et al. (1998) in the oral only condition of their validation study for the story retelling paradigm.

3.13 Transcription, Coding and Segmenting of Discourse Samples

The following section describes the procedures used in preparing the printed transcripts of the discourse audio recordings (i.e., transcription, coding, and segmenting) that were subsequently used in the analyses of language and discourse performance. Herein, the individuals responsible for transcribing the audio files are referred to as transcribers.

Blinding procedures. All of the audio files were fully de-identified and relabeled by the researcher such that the transcribers were blinded to the group allocation (i.e., PD vs. control) of the participant producing the discourse sample. There were no visual cues in the samples (i.e., hypomimia) that might identify group membership. Only audio data were available to the transcribers. Prior to assigning files to transcribers, the researcher reviewed each audio file to ensure that no identifying information (i.e., either participant specific or group allocation) was present in the audio sample. No files required editing for this reason. After completion of data collection, the researcher created a master list linking the original participant numbers (i.e., with group allocation information) to the de-identified audio file numbers and to the de-identified transcript files numbers. Once created, neither the researcher nor the transcribers had access to the master list until after all of the transcription, coding, segmenting and discourse analyses were completed.

Transcription, coding and segmenting. Three trained and remunerated transcribers (research assistants in the Aging and Communication Disorders Laboratory at Western University) transcribed orthographically the audio recordings of discourse samples ($N = 152$). While the experience levels of the transcribers varied (6 months to 3

years), each of the transcribers completed a similar training program with a minimum of 10 hours formal training in transcription that included direct teaching of discourse transcription and coding procedures, mentored transcription of audio/video samples and successful completion (i.e. inter-rater ICCs with the researcher $\geq .80$) of 5 to 10 practice transcriptions prior to transcribing any of the study samples. The complete transcription process for the audio data recordings required approximately 90 minutes per discourse sample for a total of approximately 228 hours in transcription time.

The transcribers listened to the audio recordings using Audacity® 2.0.5 (<http://audacity.sourceforge.net>). The transcriptions were created and saved as Microsoft Word documents. Transcribers were provided with the following instructions:

1. Listen to the sample one time through without stopping to familiarize yourself with the sample.
2. Listen to the sample stopping and re-playing (with or without altering playback rate) as often as necessary to correctly transcribe the sample orthographically.
3. Code the sample using the Codes for Human Analyses of Transcripts (CHAT) symbols provided.
4. Segment the sample into C-units using the definitions provided.
5. After completing the orthographic transcription and coding, replay the audio recording a minimum of two times to verify your typed transcribed sample for accuracy.

To capture details in the discourse samples not accessible via orthographic transcription alone (e.g., repetitions, revisions, incomplete words), transcribers used a pre-determined minimal set of coding conventions (Appendix D) that were adapted from *The CHILDES Project Tools for Analyzing Talk – Electronic Edition Part 1: The CHAT Transcription Format* manual (MacWhinney, 2014). Transcribers segmented the orthographically transcribed discourse samples into C-units (i.e., communication units) using the Systematic Analysis of Language Transcription (SALT) published in the respective on-

line manual (SALT, LLC, n.d.) (Appendix E). A C-unit is defined as a minimal unit of language consisting of a clause (i.e., made up of a noun and a verb) and its respective modifiers. The processes of coding and segmenting are dependent on both verbal (e.g., phonemes, words) and non-verbal (e.g., intonation, pausing) features. The processes of coding and segmenting must be completed from the audio files during the transcription stage and prior to conducting discourse analyses. After a transcriber completed the process of transcribing, coding, and segmenting the discourse samples, a second ‘expert’ transcriber (i.e., naïve to the source data) verified the transcription, coding, and segmenting accuracy against the audio file and made any corrections necessary. The final product was a rigorously transcribed orthographic representation of the audio file (i.e., $N = 152$) that was of sufficiently high quality for conducting analyses of discourse. A sample of an orthographically transcribed, coded and segmented audio file is presented in Appendix F.

Reliability study. Using an on-line random number generator (i.e., Random.org, described earlier), the transcription process was repeated by the original transcriber for 10% of discourse audio files ($n = 15$; minimum one month duration between transcription events). An equal proportion of PD and control samples were included in the reliability study. The audio files in the reliability study were assigned new sample identification numbers and were blinded again such that neither the original file information nor the original transcriber information were accessible to transcribers during the reliability study. The researcher conducted a reliability study between the original and re-transcribed samples for the number of correctly transcribed words and correctly segmented C-units. For the inter-rater reliability study, a second transcriber naïve to the original transcriptions and audio files re-transcribed a randomly generated sample of 10% of audio files ($N = 15$) of the original discourse samples. Otherwise, the procedures were identical to those used in the intra-rater reliability study. Intra-rater and Inter-rater reliability were calculated using intraclass correlation coefficients because of the applicability of this statistical procedure with for reliability studies using interval level data (Rankin & Stokes, 1998).

For total words transcribed both intra-rater and inter-rater reliability were excellent (Cicchetti, 1994) with ICC (1, 1) = 1.00, 95% CI [.99, 1.00] and ICC (3, 1) = 0.97, 95% CI [.91, .99], respectively. Similarly, for total C-units both intra-rater and inter-rater reliability were interpreted as excellent (Cicchetti, 1994) with ICC (1, 1) = 0.97, 95% CI [.91, .99] and ICC (3, 1) = 0.95, 95% CI [.83, .98], respectively. Discrepancies, when they occurred, usually involved: a) transcribing a contracted form as its two constituent words, b) omission of one or more iterations of a word repetition, c) word omissions or additions of incorrect words, and d) segmenting of utterances relative to list-like productions of consecutive main clauses not marked by the conjunction *and* which was a particular issue almost exclusively noted in the PD data. Transcription discrepancies were resolved through reviewing the audio recordings and discussion between the researcher and the transcribers. Although the reliability measures differ (e.g., most researchers in the discourse literature report simple percent agreement measures) the high degree of observed inter and intra-rater reliability for transcription accuracy is consistent with previous discourse studies (Capilouto et al., 2005; Murray, 2000). Importantly, the excellent reliability observed for transcription indicates a low degree of measurement error in the data increasing the power of the statistical tests used for analysing these data and reducing risk for Type II error (Hallgren, 2012).

3.14 Discourse Analyses

Blinding Procedures. The printed files of the orthographically transcribed discourse samples (i.e., transcripts) were fully de-identified meaning that no identifying information relative to group allocation or the individual participant appeared in the file header (See Appendix F). For the purposes of statistical analyses, the transcript was linked to the original study number using a master list and the procedure described earlier in section 3.13.

Discourse Analyses. All of the discourse analyses (DA) were conducted from the transcripts. The audio recordings were not used in the DA process. The researcher, who is experienced in discourse analyses for both research and clinical purposes, conducted the DA. Analyses were conducted using a printed version of the transcript

with manual coding by the researcher (e.g., crossing out words not used in the word count, labeling words used in the CIU count). The data for each analysis, once completed, were recorded into a data collection form (see Appendix G) that was used to enter the data into an electronic spreadsheet for later use in the statistical analyses. The protocol used in this study required approximately 90 minutes per transcript for a total of 228 hours in discourse analyses. A multi-level discourse analysis approach was used which was developed from the procedures reported within the corpus of published discourse analyses methods. The analyses conducted, the respective constructs measured, and the manner in which data variables were used is presented in Table 7. Detailed definitions of each analysis conducted and the procedure/rules followed for each analysis are presented in Appendices H and I.

Table 7

Multi-level Discourse Analyses Protocol

Analysis	Construct measured	Use of variable	
<i>Microlinguistic analyses</i>		<i>Dependent variable in analyses</i>	<i>Component of a variable</i>
Total Words	<i>Productivity</i> measured in linguistic units meeting the definition for words.	X	X
Duration of discourse sample with inter-word and inter-C-unit pauses included	Total duration of the task-relevant speaking time including pauses		X
Total C-Units	Total number of minimal language unit comprised of a clause and its modifiers		X
Words per Minute (WPM)	<i>Productivity</i> of words controlling for duration of discourse sample. Higher WPM suggests increased efficiency of spoken output but not accuracy or language content/information per se.	X	
Number of words per C-unit (i.e., Words/C-unit)	<i>Productivity</i> measured in length of minimal language units. Longer units (i.e., measured in words) suggest increased language complexity.	X	
Percentage of grammatically well-formed C-units (% Grammatical)	<i>Grammar Accuracy.</i> Proportion of C-units without errors in accuracy or completeness of grammar structures. Higher percentages suggest increased grammatical accuracy.	X	
Percentage of C-units with complex grammar structures (% Complex)	<i>Grammar Complexity.</i> Proportion of C-units with complex grammar structures (i.e., dependent clauses and/or phrasal adjuncts). Higher percentages suggest increased grammatical complexity.	X	
Percentage open class words (i.e., % Open)	<i>Lexical Use.</i> Proportion of intelligible, correct, words that meet the criteria for open class words (i.e. nouns, verbs, adjectives, and most adverbs). Reflects patterns of lexical use such as telegraphic language (i.e., elevated % Open) or reduced ability to retrieve/produce content words (i.e., reduced % Open).	X	

Percentage Verbs (i.e., % Verbs)	<i>Lexical Use.</i> Proportion of intelligible, correct, words that are verbs (i.e., excluding auxiliary and modal verbs). Reflects use of ‘main’, content bearing verbs in spontaneous spoken language, which was of special interest given the population under study.	X	
Total number of verbal disruptions (i.e., Total Disruptions)	<i>Verbal Fragmentation.</i> Total number of all instances of verbal disruptions (i.e., verbal behaviours that disrupt the flow of spoken language output). Six different verbal disruption behaviours (e.g., pauses > 2 seconds, revisions) were catalogued and totaled.	X	
Number of verbal Disruptions per 100 Words (i.e., Disrupt/100 Words)	<i>Verbal Fragmentation.</i> Verbal disruption behaviours controlling for number of total words. Higher numbers reflect a larger proportion of verbal disruptions per total words.	X	
<i>Macrostructural Analyses</i>			
Correct Information Units (i.e., CIUs) (Nicholas & Brookshire, 1993)	<i>Informativeness.</i> Intelligible words (i.e., independent of accuracy of grammatical use) which are accurate and relevant to the stimulus.	X	
Percentage of CIUs per total words (i.e., % CIUs)	<i>Informativeness.</i> Proportion of total words that meet criteria for CIUs. Higher values suggest greater efficiency of and volume of information conveyed.	X	
Number of CIUs per minute (i.e., CIUs/Minute)	<i>Informativeness.</i> CIUs controlling for duration of discourse sample. Higher CIUs/minute suggest greater efficiency of information in the time domain.	X	

Reliability study. Using an on-line random number generator (i.e., Random.org, described earlier), 10% of transcripts ($N = 15$) were reanalyzed by the researcher to assess intra-rater reliability. An equal proportion of PD and control samples were included in the reliability study. The researcher conducted a reliability study between the original and reanalyzed transcripts for CIUs, % CIUs, % Grammatical, % Complex, % Open, % Verbs, Total Disruptions, and Duration. Reliability for Total Words and Total C-units was analyzed in the transcription reliability study (See section 3.13). Since the

productivity measures are largely mathematical derivatives of Total Words, Total C-units, and duration a reliability study was not specifically conducted for WPM or Words/C-unit. Using similar methods an inter-rater reliability study (i.e., a second scorer naïve to the original analysis re-analyzed the transcripts) was conducted on a second randomly generated sample comprised of 10% ($N = 15$) of the original discourse transcripts. For the inter-rater reliability study, the scorer was provided with the definitions presented in Appendices H and I along with samples of analyzed transcripts and data forms (Appendix G). The scorer used for the inter-rater study was not involved in the transcription of data files and did not have access to the audio recordings. The scorer for the inter-rater study was a second year graduate student in speech-language pathology enrolled in the School of Communication Sciences and Disorders at Western University. The scorer was oriented to the study protocol and to the definitions used for each discourse measure. However, no formal training in discourse analysis was provided to her. The scorer had no prior research experience in discourse analysis. Intra-rater and Inter-rater reliability were analysed using ICC statistics, a robust test for reliability when interval level data are reported (Rankin & Stokes, 1998). The ICC values and qualitative interpretations (Cicchetti, 1994) for intra-rater and inter-reliability are presented in Table 8.

Table 8

Intra-rater and Inter-rater ICC Results for Discourse Measures Reliability Study

Discourse Measure	ICC (1, 1), 95% CI Intra-rater	Interpretation (Intra)	ICC (3, 1), 95% CI Inter-rater	Interpretation (Inter)
CIUs	1.00, [.99, 1.00]	Excellent	0.96, [0.79, 0.99]	Excellent
% CIUs	1.00, [.98, 1.00]	Excellent	0.82, [0.17, 0.96]	Excellent
% Grammatical	0.97, [0.84, 0.99]	Excellent	0.70, [0.05, 0.93]	Good
% Complex	0.98, [0.92, 1.00]	Excellent	0.90, [0.61, 0.98]	Excellent
% Open	0.97, [0.91, 0.99]	Excellent	0.80, [0.29, 0.96]	Excellent

% Verb	0.99, [0.95, 1.00]	Excellent	0.99, [0.96, 1.00]	Excellent
Total Disrupt	1.00, [0.98, 1.00]	Excellent	0.91, [0.62, 0.98]	Excellent
Duration	0.99, [0.98-1.00]	Excellent	0.99 [0.98, 1.00]	Excellent

The results for the inter-rater and intra-reliability were good to excellent for all discourse measures. Since a single researcher performed all of the analyses, the intra-rater reliability data are the critical values in these analyses. The high intra-rater reliability values suggest that measurement error was low in the data reducing the risk for Type II error (Hallgren, 2012). The high intra-rater reliability values demonstrate the soundness of the definitions used for the discourse analyses in that the scorer in the reliability study, while familiar with the theoretical constructs, was able to apply the constructs used with minimal practical experience and no formal training in discourse analyses other than an introduction to the procedure and the definitions.

3.15 Statistical Analyses

General statistical method. All statistical tests were conducted using IBM SPSS Version 22.0 (IBM, 2013). With the exception of sex, handedness, and primary language, the descriptive data for all variables under study are reported using means, standard deviations, and 95% confidence intervals.

Assumptions testing. The statistical analyses relied on five different parametric statistical tests: (a) independent groups t-test, (b) univariate *F* tests specifically one-way ANOVA, (c) multivariate *F* tests specifically a mixed MANOVA (between and within subject effects) with subsequent univariate tests, (d) discriminant function analysis, and (e) Pearson product moment correlations. To ensure that the data were sufficiently robust for conducting these analyses the assumptions testing was based on the most stringent test; the mixed MANOVA. The discriminant function analysis is generally robust when the assumptions of MANOVA are satisfied. The data were evaluated for independent observations, normality (univariate and multivariate), homogeneity of variance, homoscedasticity where appropriate, linearity/multicollinearity, missing data, and outliers

(univariate and multivariate) (Tabachnick & Fidell, 2006). The details of this testing and the decisions made relative to the selection of statistical procedures are outlined as follows.

Testing independent observations. In these data there were no reasons to suspect that the observations were not independent. However, the assumption was formally tested by regressing each of the individual dependent variables against the participant number that was assigned consecutively in the order that data were collected. There were no significant findings from the regression analysis therefore, independence of observations was assumed.

Testing normality. Both univariate and multivariate normality were assessed using visual inspection of the data, Skewness/ SEM_{Skew} and Kurtosis/ $SEM_{Kurtosis}$ scores, and the Shapiro-Wilk statistic. For multivariate normality the dependent variables for the discourse measures were evaluated in every possible combination of IVs x DVs. Variables with either a Skewness/ SEM_{Skew} or Kurtosis/ $SEM_{Kurtosis}$ value between -2 and +2 were considered to have a normal distribution. Given the conservative nature of the Shapiro-Wilk test a significance level of .01 was applied to the testing of all variables.

With equal n sizes in each group and $df > 20$ for all statistical tests, it was determined that both the univariate and multivariate tests used in the statistical analyses would be sufficiently robust against the identified violations of normality (Mardia, 1971). T-tests and one-way ANOVA procedures also are quite robust to violations of normality when n sizes are equal (Khan & Rayner, 2003; Mardia, 1971).

Testing homogeneity of variance, homoscedasticity, and sphericity.

Homogeneity of variance in the univariate F tests was assessed using Levene's test with a significance level of .05. Homoscedasticity for the covariance matrices for the multivariate F test was assessed using Box's M with a significance level of 0.001 given the N size for this study (i.e., < 50). Violations in homogeneity and homoscedasticity were observed for some variables such that the null hypothesis for homogeneity of variance and homoscedasticity assumptions was not rejected for all analyses. However, the robustness of both the one-way ANOVA and MANOVA procedures to these

violations given the presence of equal n sizes in the current study did not preclude the use of these statistical tests despite the violation of this assumption for some variables (Hakstian, Roed, & Lind, 1979). As a conservative approach, whenever violations in homogeneity of variance occurred in the univariate tests, the robustness of the findings was validated using Brown-Forsythe's F . If the Brown-Forsythe's F did not result in a change of the interpretation of the results then the unadjusted F values from the original statistics for the uncorrected df were reported. The multivariate effects of each MANOVA were reported as the more conservative Pillai's Trace statistic because violations in homoscedasticity were observed. In cases of violations of homogeneity of variance in the t-tests, the statistics for the corrected df were reported. In repeated measures MANOVA, Mauchly's test is typically used to assess sphericity. However, because there was only a single pair of comparisons (Picture vs. Retell) a Mauchly's test statistic was not produced by IBM SPSS Version 22. Given the ubiquitous presence of sphericity violations in repeated measures procedures, the data were assumed to violate sphericity and the Greenhouse-Geisser corrected values were reported for all within-subjects tests.

Testing the linear relationship between DVs. The linear relationship between each combination of dependent variables for each multivariate analysis was assessed visually using scatterplots and regression lines generated in IBM SPSS. Using visual analysis there were no major concerns regarding linearity for any of the canonical variable groupings.

Missing data. In the initial data set, there were 2 missing data points out of 1290 cells of data. Both were on the Trail Making Test. One PD participant was unable to complete the assessment because of severe task-specific tremors associated with writing. One control participant refused to complete the task because of a hand injury. Both data points were imputed using the average group value for that measure, a procedure that is reasonably well tolerated within missing data computations in MANOVA (Schafer & Graham, 2002). In the final data used for analyses there were no missing data points/empty cells. The absence of missing data strengthens the robustness of the MANOVA procedures used in the reported statistical analyses.

Outliers. The data were assessed for the presence of univariate and multivariate outliers for all dependent variables, both individual (DiVs) and canonical (DcVs). Each data point was converted to a z score to assess for univariate outliers. Based on the n size (i.e., < 50) values ± 2.5 SDs from the DiV mean were flagged for review. Each of the flagged data points was manually reviewed for potential data entry errors, abnormal participant characteristics that may have affected the data, a consistent source of outlier data, and analysis errors. Outliers resulting from data entry or data analysis errors were corrected and the univariate outlier assessment procedure repeated. Remaining outliers were again flagged and reviewed. The outliers were not removed from analysis because none of the remaining outliers were associated with a particular source or atypical participant characteristics (e.g., older age, lower education, lower cognitive scores). Overall, a total of 4 univariate outliers ($< 1\%$ of total data points) remained in the cognitive and language test data and 18 univariate outliers ($< 1\%$ of total data points) remained in the discourse data. The presence of multivariate outliers was assessed using Mahalanobis distances. A Mahalanobis distance for each data point in each of five canonical variables (i.e., discourse measures) was derived and then evaluated for statistical significance using a Chi squared statistic ($\alpha = 0.001$) and the df specific to the canonical variable. Once the individual univariate outliers related to data entry were corrected, there were no multivariate outliers for any of the canonical variable groupings.

Statistical tests procedures. After completing the testing of assumptions, it was determined that the data were sufficiently robust to use the proposed parametric tests. A detailed description is outlined below for the statistical procedures used to analyze each data set and to answer the research questions.

Demographic variables. A series of 3 separate independent t-tests with Group as the IV and age, years of formal education and Mattis Dementia Rating Scale -2 (MDRS-2) as respective DVs were performed. The series was conducted to determine whether PD and controls differed significantly on any of the demographic variables. For ease of interpretation by the reader MDRS-2 scores were reported both as an unadjusted total test score and as MOANS age-adjusted scaled scores. As recommended, the MOANS scaled score was used in the statistical test of group differences because age has a significant

influence on MDRS-2 scores (Strauss, Sherman, & Spreen, 2006). All other variables including sex, UPDRS III scores, Hoehn and Yahr scores, duration of PD, and levodopa equivalent dose (LED) were reported using descriptive methods of central tendency.

Analyzing the cognitive and speech intelligibility/rate measures. A multivariate analysis approach did not conceptually fit with the cognitive and language measures conducted as a part of the protocol. Therefore, ten separate one-way ANOVAs (with Group as the IV and cognitive/language test score as the DV) corrected for multiple comparisons bias were conducted to evaluate whether or not the groups significantly differed in terms of performance on a battery of standardized cognitive and language measures. A significance level of .10 was established then corrected such that each test was evaluated against a significance criterion of .01. Two separate one-way univariate ANOVAs were conducted using a Bonferroni corrected significance level of .05 to elucidate potential speech intelligibility and speech rate differences between groups. Scores from the SIT and speech rate in words per minute were analyzed as DVs with Group (PD vs. control) as the independent variable for both ANOVAs.

Analyzing the discourse measures and determining the minimal set of discriminative discourse variables. Two stimuli for each discourse genre (Picture and Retell) were collected to optimize sampling. The stimuli used for each genre are well established in the field of discourse research and have been validated previously as parallel forms (Brookshire & Nicholas, 1994b; Capilouto et al., 2005; Doyle et al., 2000). As such, they were assumed to produce equivalent discourse samples. For each variable, the discourse data for the two parallel form stimuli for each genre were averaged such that a single score was entered for Picture and a single score for Retell.

An a priori grouping of twelve discourse variables (DiVs) into five separate canonical variables (DcVs) was completed. The assignment of the discourse variables into canonical groupings representing domains of discourse impairment was informed theoretically by the existing literature on discourse performance. The canonical variables were:

- DcV 1: Productivity comprised of Total Words, Words/C-unit, WPM;

- DcV 2: Grammar comprised of % Grammatical utterances and % Complex;
- DcV 3: Lexical comprised of % Open class words and % Verbs;
- DcV 4: Informativeness comprised of Total CIUs, % CIUs, and CIUs per minute;
- DcV 5: Verbal Fragmentation comprised of Total verbal disruptions, Disrupt/100 Words.

Using the canonical variables, five separate mixed MANOVAs (one for each DcV group) were conducted with two IVs: Group (PD vs. Control) and Task (Picture vs. Retell). Group was treated as a between subjects IV while Task was treated as a within subjects IV. Pillai's Trace statistic was reported for all multivariate effects. MANOVAs and subsequent univariate *F* tests were conducted at a significance level of .10 for reasons discussed at length in section 3.1 under the subheading *Sample size determination*. Of primary interest for the objective of the current study were the main effects of Group for all analyses. However, the effects of Task and Task by Group interactions were of paramount importance to the overall protocol given that the researcher intentionally sampled across two different tasks to increase the representativeness of the discourse sample relative to spontaneous spoken language. Therefore, prior to analyzing the discourse data as a single representative sample, independent of task, it was important to evaluate Group x Task interactions that could, pending the source of the interaction, affect the further analysis/interpretation of group effects. A series of univariate *F* tests for the individual DiVs within a canonical variable were conducted subsequent to the MANOVA procedures for both Group and Task effects. The main effects of Group were used to determine which individual DiVs contributed significantly to the separation of groups to be used in the predictive discriminant function analysis. While all data and statistical tests were reported for the purposes of transparency and comprehensiveness, only those ANOVAs subsequent to significant multivariate tests were interpreted in detail.

Analyzing the optimized weighting of variables for discriminating the discourse of PD vs. controls. Discriminant function analysis (DFA) was used to determine the optimized combination of discourse variables (DiVs) for discriminating the discourse performance of participants with PD vs. controls. The variables entered as predictors in

the DFA were derived from the series of MANOVAs and subsequent univariate F tests that assessed for the effect of group on measures of discourse performance. The discriminant function analysis was conducted using IBM SPSS Version 22.0. Group was entered as the DV with the IV's derived from the significant DiVs identified in the post-MANOVA univariate analyses. For the DFA, the discriminant variable was cross-validated via an external analysis (i.e., applying the classification developed on one data set to a different data set) using the leave-one-out analysis option in IBM SPSS Version 22.0 package. Prior probabilities were computed from the 'group sizes' option in IBM SPSS because the existing literature in PD was not sufficient for determining prior probability information. The classification function used in the DFA was the Classification Function Coefficient option in IBM SPSS Version 22.0. At the conclusion of the analyses, a unique discriminant function score for each discourse sample was calculated ($N = 76$). The individual scores were plotted graphically to assess visually if the accuracy of the discriminant function in predicting group membership differed as a function of Task (Picture vs. Retell).

Analyzing the relationship between disease variables, demographic variables and discourse performance in PD. Four separate bivariate Pearson's Product Moment correlations were calculated to determine the relationships among disease variables, relevant demographic variables and discourse performance. Results were evaluated at significance levels of .01 and .05. Separate bivariate correlations were conducted to assess the following associations:

- Individual discriminant function scores for Retell with UPDRS-III, Hoehn and Yahr, Duration, and LED;
- Individual discriminant function scores for Picture with UPDRS-III, Hoehn and Yahr, Duration, and LED;
- Individual discriminant function scores for Retell tasks with Age, Education, and MDRS-2 scores; and
- Individual discriminant function scores for Picture tasks with Age, Education, and MDRS-2 scores.

Significance levels and multiple comparisons bias corrections. Unless otherwise noted a significance of .10 was used for all MANOVA/ANOVA procedures. T-tests were conducted with a significance level of .05 because of their ability to maintain sufficient power in the current design with the more stringent significance levels. Statistically significant Pearson's Product-Moment correlations were considered relative to .05 and .01 levels of significance. Whenever multiple individual one-way ANOVAs were used to analyze the effect of an IV on separate DVs (e.g., speech intelligibility, BNT), significance levels were adjusted for multiple comparisons bias using a Bonferroni correction procedure applied to an initial significance level of .10 (for reasons discussed in section 3.1). However, univariate F tests used to assess the effects associated with individual DVs conducted in follow-up to a significant multivariate F test for the canonical variable were assessed at an uncorrected significance level of 0.10. Hummell and Sligo (1971) proposed that when a multivariate F test procedure is significant the subsequent univariate tests are protected from multiple comparisons bias. Therefore, a Bonferroni correction was not applied to any of the follow-up univariate between groups tests performed in conjunction with a MANOVA.

Effect size reporting. Unless otherwise noted, effect sizes for both the multivariate and subsequent individual univariate F tests were reported as η^2_p (partial eta squared). Partial eta squared measures the proportion of variance in the variable under study explained by the individual effect of a target variable after controlling for the variance associated with the remaining variables. In a one-way MANOVA the values for eta squared and partial eta squared are typically equal; however, the interpretation between the two differs. Partial eta squared was preferred rather than Eta squared because the former enables variance comparisons across canonical variables. Partial eta squared was reported from the IBM SPSS Version 22.0 output. The following effect size conventions were applied for the qualitative interpretation qualitatively partial eta squared values relative to effect size: small = .01, medium = .09, and large = .25 (Cohen, 1988). For t-tests, Cohen's d was used as the measure of effect size and interpreted using the following conventions: small = 0.20, medium = 0.50, and large = 0.80 (Cohen, 1988).

Chapter 4: Results

The results of the statistical analyses are presented in the following sections. The Results chapter follows the order of the research questions as they are presented at the conclusion of Chapter 2.

4.1 Cognition, Language and Speech Intelligibility Measures (RQ 1)

Descriptive data and statistical test results for the cognitive and language measures are reported in Table 9. The DV distributions were not platykurtic with kurtosis estimates for all variables falling above -1.0. There were violations of normality on the more conservative Shapiro-Wilk's test ($\alpha = .01$) for the following variables including: (a) control Verb naming: TAWF ($p < .000$); (b) control The Pyramids and Palm Trees test ($p = .001$); (c) PD semantic fluency ($p = .004$); PD SPPT ($p < .001$); and (d) control SPPT ($p < .001$). Variables with significant Shapiro-Wilk's tests that did not violate the secondary criteria for normality using Skewness/ SEM_{Skew} and Kurtosis/ $SEM_{Kurtosis}$ values include control Verb naming (i.e., TAWF-verb) and control The Pyramids and Palm Trees Test. For the remaining variables that violated normality, Skewness/ SEM_{Skew} values ranged from 2.02 to 2.30 with the Kurtosis/ $SEM_{Kurtosis}$ values < 2 . These values are in acceptable limits for violations of skewness and kurtosis (Tabachnick and Fidell, 2006). Given the robustness of ANOVA procedures to violations of non-normality when n sizes are equal these values do not substantially affect the interpretation of the results (Khan & Rayner, 2003).

Levene's test ($\alpha = .05$) were significant for (a) Trail Making Test ($F(1, 36) = 7.31, p = .010$); (b) Verb naming: TAWF ($F(1, 36) = 19.10, p < .001$); (c) The Pyramids and Palm Trees Test ($F(1, 36) = 6.13, p = .018$); and (d) SPPT ($F(1, 36) = 8.59, p = .049$).

Levene's tests were non-significant for (a) Stroop ($F(1, 36) = 0.02, p = .888$); (b) ABCD verbal learning ($F(1, 36) = 0.002, p = .960$); (c) Phonemic fluency ($F(1, 36) = 0.10, p = .757$); (e) Action fluency ($F(1, 36) = 0.19, p = .667$); (f) Semantic fluency ($F(1, 36) = 0.04, p = .849$); and (g) BNT ($F(1, 36) = 0.41, p = .528$). While violations in univariate homogeneity of variance for between group comparisons were observed for some data

sets, Brown–Forsythe F adjustments showed this had no impact on the observed outcomes. Therefore, the F values for the uncorrected degrees of freedom are reported.

Table 9

Performance on Cognitive and Language Measures for participants with PD and Controls (Ctrl)

Dependent Variable	df	df Error	F	p	ES	Group	Mean (SD)	Lower Bound	Upper Bound
Trail Making Test (B/A ratio)	1	36	1.37	.249	<i>n.s.</i>	PD	2.56(1.04)	2.06	3.06
						Ctrl	2.25(.47)	2.03	2.48
Stroop (Interference)	1	36	0.02	.891	<i>n.s.</i>	PD	.93(.26)	0.81	1.06
						Ctrl	.94(.23)	0.84	1.05
ABCD verbal learning	1	36	2.74	.107	<i>n.s.</i>	PD	6.58(2.27)	5.49	7.67
						Ctrl	7.84(2.43)	6.67	9.01
Phonemic Fluency	1	36	2.44	.127	<i>n.s.</i>	PD	36.12(12.55)	30.07	42.17
						Ctrl	42.42(12.30)	36.49	48.35
Action Fluency	1	36	8.17	.007	.159	PD	13.63(6.75)	10.38	16.89
						Ctrl	19.32(5.44)	16.70	21.94
Semantic Fluency	1	36	3.21	.081	<i>n.s.</i>	PD	17.17(7.03)	13.77	20.55
						Ctrl	21.26(7.09)	17.88	24.68

Boston Naming Test	1	36	3.26	.080	<i>n.s.</i>	PD	26.47(2.34)	25.34	27.60
						Ctrl	27.79(2.15)	26.75	28.83
TAWF-verb	1	36	15.71	< .001	.279	PD	23.32(1.29)	22.69	23.94
						Ctrl	24.58(.51)	24.33	24.82
The Pyramids and Palm Trees Test	1	36	4.27	.046	<i>n.s.</i>	PD	50.47(1.61)	49.70	51.25
						Ctrl	51.32(.75)	50.95	51.68
Sentence Production Priming Test	1	36	2.31	.138	<i>n.s.</i>	PD	28.95(1.47)	28.24	29.66
						Ctrl	29.53(.77)	29.15	29.90

Note. Significance level for all tests reported in the table = .01. Effect sizes (ES) are reported as partial eta squared values. The lower and upper bounds were calculated for 95% confidence intervals. Non-significant statistical tests are denoted by *n.s.* in the column labelled ES. Significant *p* values are bolded.

The results indicate that participants with PD differed significantly from controls on two measures assessing word retrieval for verbs including TAWF-verb and action fluency. Large effect sizes were observed for verb confrontation naming (Cohen, 1988). Medium to large effect sizes were observed for action fluency (Cohen, 1988). These are important results in the context of increasing focus on verb-specific impairments in the PD literature (Bertella et al., 2002). There were no significant differences for traditional measures of frontal lobe functions examining attention, switching, reaction time, and cognitive flexibility processes. Furthermore, PD and controls performed similarly on a coterie of language measures including confrontation naming for nouns, semantic knowledge for objects, word retrieval for semantic categories, and syntax production.

Speech intelligibility and rate in relation to Group. It was important to evaluate PD vs. control differences in motor speech performance given the crucial aspect of speech intelligibility for spontaneous language tasks and the ubiquitous presence of motor speech changes in PD (Ramig et al., 2004). Group differences for the Sentence

Intelligibility Test (SIT) and speech rate (i.e., words per minute calculated from the Grandfather Passage) were analyzed. A significant Shapiro-Wilk's test (significance set at .01) was found for the SIT data for both the PD group ($p < .001$) and the control group ($p < .001$). However, the Skewness/ SEM_{Skewness} and Kurtosis/ SEM_{Kurtosis} ratios did not exceed acceptable limits. In contrast the data for baseline speech rates were distributed normally for both PD (Shapiro-Wilk's $p = .678$) and controls (Shapiro-Wilk's $p = .865$). The violations were not deemed a substantial concern given the robustness of ANOVAs to violations of normality when n sizes are equal (Khan & Rayner, 2003). Levene's test was not significant for either the SIT ($F(1, 36) = 1.80, p = .186$) or baseline speech rate ($F = 1.78, p = .191$) with a significance of .05.

There was no significant difference in speech intelligibility on the SIT between PD ($M = 94.51, SD = 9.55, 95\% \text{ CI } [89.91, 99.11]$) vs. controls ($M = 97.51, SD = 2.77, 95\% \text{ CI } [96.18, 98.85]$) $F(1, 36) = 1.73, p = .191$. This result indicates that speech intelligibility did not differ significantly between PD and controls. It was important to examine baseline speech rates between groups because multiple discourse measures were reported as rate measures (e.g., CIUs/Minute). The results showed no significant effect of group on baseline speech rates $F(1, 36) = 1.87, p = .180$. Speech rates at baseline did not differ significantly between PD ($M = 142.88, SD = 27.21, 95\% \text{ CI } [130.65, 155.12]$) vs. controls ($M = 153.13, SD = 18.10, 95\% \text{ CI } [145.0, 161.27]$).

4.2 Discourse Measures in Relation to Group x Task and Task Effects (RQ 2)

One of the unique features of the current study was the use of a cross-genre sampling method to obtain a representative sample of spontaneous language. The a priori assumption, based on the existing discourse literature for healthy adults and persons with aphasia, was that sampling tasks from two genres that differed in their inherent cognitive-linguistic demands would increase the representativeness and robustness of the discourse data (Brookshire & Nicholas, 1994b; Brookshire & Nicholas, 1994a; McNeil, Doyle, Fossett, Park, & Goda, 2001; McNeil et al., 2007). Optimizing sample representativeness was key for reducing a large array of discourse measures to a minimal set of variables that would distinguish optimally the discourse of participants with PD vs. controls.

However, because the discourse protocol used in the current study was not reported previously in PD, prior to interpreting the group effects for this purpose it was necessary to examine the possibility of Group by Task interactions for each of the canonical variables. The presence of interactions could affect the integrity of the assumption that the tasks added useful variance to the data in a consistent way for each group. Furthermore, to assess the added benefit of using a cross-genre vs. a single genre paradigm it was necessary to assess the presence and nature of task differences (independent of group) for each of the canonical variables. Five separate mixed methods MANOVAs were conducted, as detailed in the Method section, to assess for the interaction between Task and Group in order to understand better the effects in a systematic way. The subsequent multivariate and univariate tests for Task and Group were evaluated separately with the analysis and interpretation of the results focusing on the Group effects as they pertained to the primary objective of developing a minimal set of discriminative discourse variables.

Each of the separate mixed MANOVAs was run with two IVs: Group as the between-subjects variable and Task as the within-subjects variable based on a significance level of .10. The individual discourse measures specific to each of five canonical variables (DiVs) were entered as DcVs: DcV 1: Productivity (DiVs = 3); DcV 2: Grammar (DiVs = 2); DcV 3: Lexical (DiVs = 2); DcV 4: Informativeness (DiVs = 3); and DcV 5: Verbal Disruptions (DiVs = 3). The individual univariate tests were interpreted at a significance level of .10.

Two DiVs had a significant Shapiro-Wilk's test: PD Total Words ($p = .003$) with a Skewness/ SEM_{Skew} value of 2.07; and control Words/C-unit ($p = .001$) with a Skewness/ SEM_{Skew} value of 2.2 and a Kurtosis/ $SEM_{Kurtosis}$ value 2.11. These variables were left in the analysis because they were judged to be a low threat to normality. Since none of the data were platykurtic and the MANOVA procedure is relatively robust to minor violations in multivariate normality when n sizes are equal and $df > 20$, the observed violations of normality were not of significant concern and the variables were not removed from the analyses (Mardia, 1971).

With a significance criteria of $< .001$, the tests of homogeneity of covariance were significant for one of the five canonical variables (value bolded):

- Productivity, Box's $M = 32.18$ was associated with a p value of .194;
- Grammar, Box's $M = 37.33$ was associated with a p value of **$< .001$** ;
- Lexical, Box's $M = 5.08$ was associated with a p value of .924;
- Informativeness, Box's $M = 51.14$ was associated with a p value of .004;
- Verbal Fragmentation, Box's $M = 33.18$ was associated with a p value of .001.

The violation of homoscedasticity is not of substantial concern for interpreting the results given the robustness of MANOVA to such violations when n sizes are equal between groups (Hakstian et al., 1979). For the subsequent univariate tests, Levene's test was significant for 5 of the 24 tests. A summary of the Levene's statistics results is presented in Table 10.

Table 10.

Summary of Levene's Tests for all Discourse Variables

Variable	Levene's Test Result	Variable	Levene's Test Result
Total Words Picture	$F(1, 36) = 1.78, p = .191$	% Verb Retell	$F(1, 36) = 0.03, p = .862$
Total Words Retell	$F(1, 36) = 1.28, p = .265$	Total Disruptions Picture	$F(1, 36) = 0.30, p = .585$
Words/C-unit Picture	$F(1, 36) = 0.20, p = .888$	Total Disruptions Retell	$F(1, 36) = 0.07, p = .800$
Words/C-units Retell	$F(1, 36) = 0.01, p = .937$	Disrupt/100 Words Picture	$F(1, 36) = 5.52, p = .024$
WPM Picture	$F(1, 36) = 9.82, p = .003$	Disrupt/100 Words Retell	$F(1, 36) = 0.56, p = .458$
WPM Retell	$F(1, 36) = 4.16, p = .049$	Total CIUs Picture	$F(1, 36) = 0.62, p = .437$
% Grammatical Picture	$F(1, 36) = 13.93, p = .001$	Total CIUs Retell	$F(1, 36) = 2.34, p = .135$

% Grammatical Retell	$F(1, 36) = 3.82, p = .058$	% CIUs Picture	$F(1, 36) = 5.74, p = .022$
% Complex Picture	$F(1, 36) = 3.87, p = .057$	% CIUs Retell	$F(1, 36) = 1.67, p = .205$
% Complex Retell	$F(1, 36) = 0.07, p = .795$	CIUs/Minute Picture	$F(1, 36) = 2.56, p = .119$
% Open Picture	$F(1, 36) = 0.67, p = .418$	CIUs/Minute Retell	$F(1, 36) = 1.32, p = .258$
% Open Retell	$F(1, 36) = 0.56, p = .458$		
% Verb Picture	$F(1, 36) = 0.16, p = .693$		

Note. $\alpha = .05$ for all Levene's tests. Significant values are bolded.

While there were minor violations in univariate homogeneity of variance for some of the individual discourse variables (Table 10) the Brown–Forsythe F adjustments indicated that these violations had no impact on the observed outcomes for any variable. As a result, the F values were reported for the unadjusted degrees of freedom.

A summary of the descriptive data by task, the MANOVA within-group results, and subsequent univariate test results for Task are presented in Table 11. The Task x Group multivariate effects for each mixed MANOVA were used to determine if there were any significant effects of canonical variables as a function of task between the PD and control groups. For Productivity there was a significant effect of Task (Table 11) but no Task x Group interaction ($F(3, 34) = 0.98, p = .416$). For Grammar there was no effect of the canonical variable in relation to task (Table 11) but there was a significant interaction between Task and Group ($F(2, 35) = 2.95, p = .07$). There was a significant effect of Lexical in relation to task (Table 11) but no interaction effect between Group and Task ($F(2, 35) = 0.49, p = .615$). The effect for Informativeness was significant in relation to Task (Table 11) but there was no Task by Group interaction ($F(3, 34) = 1.28, p = .297$). There was a significant effect of verbal fragmentation in relation to Task (Table 11) but no interaction between Task and Group ($F(2, 35) = 1.38, p = .266$). These results suggest that for 4 of the 5 canonical variables, there was no significant effect of the variable as a function of task between the PD and control groups.

Discourse performances for four of the five canonical variables differed in relation to Task. Moreover, the effect of the canonical variable for Task was not dependent on Group for these four variables (i.e., Productivity, Lexical, Informativeness, and Verbal Fragmentation). In other words the effects of the intrinsic demands of each task affected discourse performance in an equivalent way for both participants with PD and controls for select variables. However, for the canonical variable Grammar, with no significant task effect, there was a significant Group x Task interaction. With concerns over Group x Task interactions examined, the analyses proceeded to assess the univariate effects of Task in order to determine to what degree cross-genre sampling increased the representativeness of the discourse sample.

Table 11.

Discourse Performance Effects in Relation to Task Independent of Group

Dependent Variable	<i>df</i>	<i>df</i> Error	<i>F</i>	<i>p</i>	ES	Task	Mean (<i>SD</i>)	Lower bound	Upper bound
DeVs and DiVs									
DeV 1: Productivity	3	34	12.43	< .000	.523				
Total Words	1	36	23.82	< .000	.398	Picture	129.37(81.17)	103.56	155.18
						Retell	183.32(63.73)	163.06	203.58
Words/C-unit	1	36	10.51	.003	.226	Picture	7.84(1.49)	7.37	8.31
						Retell	8.58(1.36)	8.15	9.01
Words/Minute (WPM)	1	36	2.77	.105	<i>n.s.</i>	Picture	147.12(33.29)	136.54	157.70
						Retell	140.20(35.08)	140.20	151.35
DeV 2: Grammar	2	35	.888	.420	<i>n.s.</i>				
% Grammatical	1	36	1.83	.185	<i>n.s.</i>	Picture	90.79(8.73)	88.01	93.57
						Retell	88.33(10.76)	84.91	91.75

% Complex	1	36	.03	.865	<i>n.s.</i>	Picture	20.69(9.27)	17.74	23.64
						Retell	20.99(8.11)	18.41	23.57
DeV 3: Lexical	2	35	61.43	< .000	.778				
% Open Class Words	1	36	118.47	< .000	.767	Picture	51.43(4.52)	49.99	52.87
						Retell	44.11(3.74)	42.92	45.30
% Verbs	1	72	0.04	.843	<i>n.s.</i>	Picture	19.81(2.49)	19.02	20.60
						Retell	19.89(2.00)	19.25	20.53
DeV 4: Informativeness	3	34	12.12	< .000	.517				
CIUs	1	36	11.78	.002	.246	Picture	95.79(59.84)	76.76	114.82
						Retell	124.75(37.30)	112.89	136.61
%CIUs	1	36	15.41	< .000	.300	Picture	76.55(13.33)	72.31	80.79
						Retell	69.39(10.29)	66.12	72.66
CIUs/Minute	1	72	14.99	< .000	.294	Picture	113.29(34.06)	102.46	124.12
						Retell	98.53(30.37)	88.87	108.19
DeV 5: Verbal Fragmentation	2	35	24.73	< .000	.659				
Total Disruptions	1	36	36.80	< .000	.506	Picture	13.04(7.62)	10.62	15.46
						Retell	19.45(11.77)	15.71	23.19
Disrupt/100 Words	1	36	.417	.523	<i>n.s.</i>	Picture	13.04(7.62)	10.62	15.46
						Retell	12.21(8.17)	9.61	14.81

Note. Significance level for all tests reported in the table = .10. Pillai's Trace statistic is reported for all multivariate tests. Greenhouse-Geisser statistic is reported for all of the univariate effects. Effect sizes (ES) are reported as partial eta squared (η^2_p). The lower

and upper bounds were calculated for 95% confidence intervals. Non-significant statistical tests are denoted by *n.s.* in the column labelled Effect Size (ES).

A summary of the proportion of variance associated with Task differences attributed to each of the significant canonical variables is presented in Figure 5 to show relative contributions of variance in relation to Task.

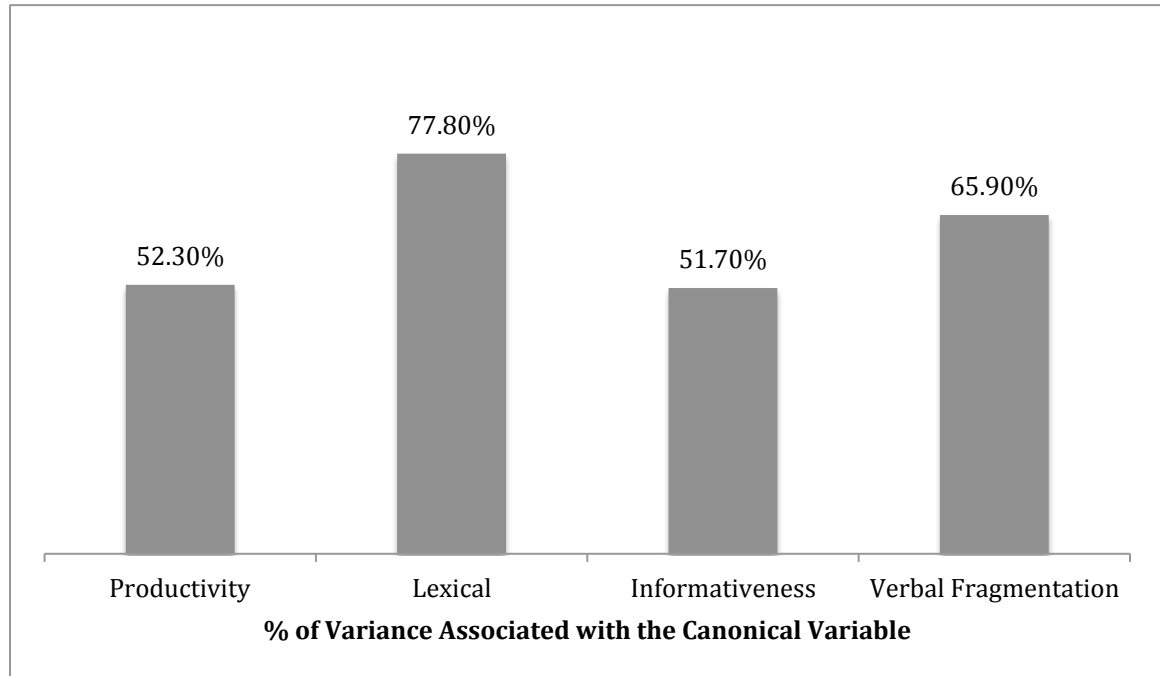


Figure 5. Summary of variance associated with each statistically significant canonical variable in relation to Task. Variances correspond to partial eta squared values and are the proportion of variance for Task accounted for by the canonical variable after controlling for all other sources of variance.

Productivity in relation to Task. There was a significant effect of the canonical variable Productivity, which is comprised of Total Words, Words/C-unit, and WPM for distinguishing tasks (Table 11). A large effect size for the multivariate effect was observed for Productivity ($\eta^2_p = .523$) (Cohen, 1988). The proportion of Task variance associated with the canonical variable Productivity is shown in Figure 5. There was a significant univariate effect of Task for Total Words and Words/C-unit, but not for WPM (Table 11). The partial eta squared values reported suggest that the effect of Task for

Total Words was of large magnitude and for Words/C-unit was of medium to large magnitude (Cohen, 1988). Participants produced significantly more total words and more words per C-unit for the Retell task vs. the Picture task (see descriptive data in Table 11).

Grammar in relation to Task. Because of the significant interaction effect, the non-significant multivariate effect for Grammar in relation to Task (Table 11) was further examined with follow-up MANOVAs conducted for the canonical variable Grammar separately for each group (i.e., PD and control). The results indicated that the effect of Task on the variable Grammar was significant for the control group $F(2, 35) = 6.63, p = .004, \eta^2_p = .275$ but not for the PD group $F(2, 35) = 1.00, p = .378$. Specifically, for the control group, the task effect was significant for the DiV % Grammatical $F(1, 36) = 10.14, p = .003, \eta^2_p = .220$ but not for % Complex $F(1, 36) = 1.03, p = .317$. For the PD group, neither % Grammatical ($F(1, 36) = .001, p = .976$) nor % Complex ($F(1, 36) = 2.06, p = .160$) differed significantly between tasks. Specifically, these data suggest that controls produced more grammatical errors with a lower proportion of well-formed sentences on the Retell task ($M = 0.91, SD = .06$) than on the Picture task ($M = 0.96, SD = .03$).

Lexical in relation to Task. There was a significant effect for the canonical variable Lexical comprised of % Open and % Verbs in relation to Task (Table 11). A partial eta squared value of .503 suggests a large multivariate effect of the variable Lexical in relation to Task (Cohen, 1988). The proportion of Task variance associated with the canonical variable Lexical is shown in Figure 5. Of the univariate tests, only % Open (i.e., proportion of total words that were open class words) was significantly different between tasks. The proportion of verbs produced was not significantly different between Picture vs. Retell tasks. The effect of task on % Open was interpreted as large ($\eta^2_p = .767$) (Cohen, 1988). As reported in Table 11, the proportion of open class words was significantly higher for the Picture task vs. the Retell task.

Informativeness in relation to Task. There was a significant effect of Informativeness in relation Task with medium to large effects observed ($\eta^2_p = .517$) (Table 11) (Cohen, 1988). The proportion of Task variance associated with the canonical

variable Informativeness is shown in Figure 5. The univariate tests for Task were interpreted subsequent to a significant multivariate effect. All three univariate variables were significant in relation to Task: Total CIUs, % CIUs, and CIUs/Minute (Table 11). The analyses showed large effect sizes (η^2_p) ranging from .246 to .300 (Cohen, 1988). However, the effect of Informativeness in relation to Task was complex. As the descriptive data in Table 11 show, the total number of CIUs was significantly higher for the Retell task vs. Picture task. In contrast, measures reflecting both proportion of content-conveying words (% CIUs) and efficiency of communicating task specific content (CIUs/Minute) were significantly higher for the Picture task vs. the Retell task.

Verbal Fragmentation in relation to Task. There was a significant effect of the canonical variable Verbal Fragmentation comprised of Total Disrupt and Disrupt/100 Words in relation to task (Table 11). The effect of Verbal Fragmentation in relation to task ($\eta^2_p = .659$) was interpreted as large (Cohen, 1988). The proportion of Task variance associated with the canonical variable Verbal Fragmentation is shown in Figure 5. In the subsequent univariate tests only Total Disruptions differed significantly between tasks with a medium sized effect (Table 11)(Cohen, 1988). In contrast, the task effect for Disrupt/100 Words was not significant. As indicated by the descriptive data presented in Table 11, the number of total disruptions was significantly higher for the Retell task vs. the Picture task.

Collectively, these results confirm that four of the five canonical variables differed in relation to the intrinsic demands of the discourse tasks. The results suggest that while Grammar remained stable across tasks, Productivity, Lexical, Informativeness, and Verbal Fragmentation all differed significantly as a function of Task. Subsequent univariate analyses showed that several of the individual discourse variables differed significantly between the Picture and the Retell task including: Total Words, Words/C-unit, % Open, CIUs, % CIUs, CIUs/Minute, and Total Disrupt. Importantly, the absence of Group x Task interactions for all of the significant canonical variables indicates that optimizing the representativeness of the discourse sample for the current study, by using both Picture and Retell tasks, did not come with the added cost of an interaction effect.

4.3 Discourse Measures in Relation to Group (RQ 3)

Central to creating a discourse discriminant function for categorizing effectively the spontaneous language of PD vs. controls was the development of a minimal array of discourse variables that characterized the differences in discourse between groups. The between subjects multivariate effects of the five canonical variables and subsequent univariate F tests were used to reduce the number discourse variables to a composite set of variables that differed significantly between PD vs. controls. Descriptive data by group, the multivariate MANOVA results and subsequent univariate test results for Group effects are presented in Table 12.

Table 12.

Discourse Performance In Relation to Group Independent of Task

DiVs and DeVs	<i>df</i>	<i>df</i> Error	<i>F</i>	<i>p</i>	ES	Group	Mean (<i>SD</i>)	Lower bound	Upper bound
DeV 1: Productivity	3	34	2.74	.058	.195				
Total Words	1	36	3.87	.057	.097	PD	136.47(74.05)	112.93	160.01
						Ctrl	176.22(76.44)	151.92	200.52
Words/C-unit	1	36	3.910	.387	<i>n.s.</i>	PD	8.03(1.40)	7.58	8.48
						Ctrl	8.39(1.52)	7.91	8.87
Words/Minute (WPM)	1	36	7.74	.009	.177	PD	130.52(37.94)	118.46	142.58
						Ctrl	156.80(23.87)	149.21	164.39
DeV 2: Grammar	2	35	6.29	.005	.264				
% Grammatical	1	36	11.92	.001	.249	PD	85.63(11.56)	81.95	19.31
						Ctrl	93.48(5.44)	91.75	95.21

% Complex	1	36	0.79	.381	<i>n.s.</i>	PD	19.89(7.98)	17.35	22.43
						Ctrl	21.79(9.28)	18.84	24.74
DeV 3: Lexical	2	35	0.16	.856	<i>n.s.</i>				
% Open Class Words	1	36	0.16	.814	<i>n.s.</i>	PD	47.91(4.93)	46.34	49.48
						Ctrl	47.63(6.13)	45.68	49.58
% Verbs	1	72	0.11	.748	<i>n.s.</i>	PD	19.75(2.24)	19.04	20.46
						Ctrl	19.95(2.27)	19.23	20.67
DeV 4: Informativeness	3	34	9.30	< .001	.451				
CIUs	1	36	13.30	.001	.270	PD	88.47(39.67)	75.86	101.08
						Ctrl	132.07(53.42)	115.09	149.05
%CIUs	1	36	9.35	.004	.206	PD	68.30(13.40)	64.04	72.56
						Ctrl	77.64(9.25)	74.70	80.58
CIUs/Minute	1	36	14.06	.001	.281	PD	90.18(33.51)	79.53	100.83
						Ctrl	121.64(23.72)	114.10	129.18
DeV 5: Verbal Fragmentation	2	35	4.01	.027	.187				
Total Disruptions	1	36	.032	.860	<i>n.s.</i>	PD	15.45(10.24)	12.19	18.71
						Ctrl	15.99(11.05)	12.48	19.50
Disrupt/100 Words	1	36	3.87	.057	.097	PD	13.86(6.68)	10.35	17.37
						Ctrl	13.81(6.73)	11.67	15.95

Note. Significance level for all tests reported in the table = .10. Pillai's Trace statistic is reported for all multivariate tests. Effect sizes are reported as partial eta squared (η^2_p). The lower and upper bounds were calculated for 95% confidence intervals. Non-significant statistical tests are denoted by *n.s.* in the column labelled Effect Size (ES).

A summary of the proportion of variance associated with Group differences attributed to each of the significant canonical variables is presented in Figure 6.

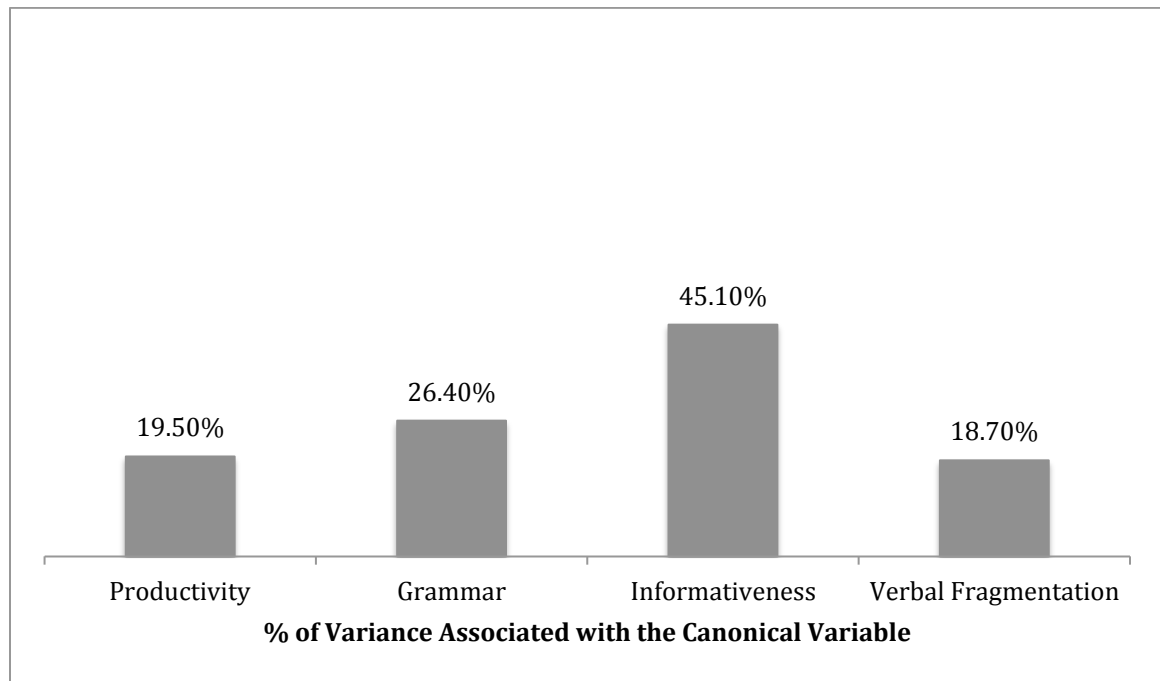


Figure 6. Summary of the proportion of variance associated with Group for each statistically significant canonical variable. Variances correspond to partial eta squared values and represent the proportion of variance explained by Group, controlling for all other variables.

Productivity in relation to Group. There was a significant effect of the canonical variable Productivity comprised of Total Words, Words/C-unit, and WPM for distinguishing groups (Table 12). The effect size ($\eta^2_p = .195$) was of medium to large magnitude using Cohen's conventions (Cohen, 1988). The proportion of variance attributable to the association between group and the canonical variable Productivity is presented in Figure 6. For the subsequent univariate tests both Total Words and WPM differed significantly between groups (Table 12). Medium sized effects were observed for both Total Words and WPM (Cohen, 1988). As the descriptive statistics in Table 12 show, Individuals with PD produced significantly fewer words and fewer words per minute vs. controls (Table 12). The following excerpt from the retell task (i.e. 'Out of

Gas') for a PD participant illustrates how verbal disruptions in the form of pauses, revisions, and incomplete utterances affected productivity in the PD group:

A man was going...[trailing off]

He was a <tave> [/] a salesman selling paint

I think

And [2 second pause] he [2 second pause] decided to go but his car broke down

So [8 second pause] he went on ... [trailing off]

Grammar in relation to Group. There was a significant effect of the canonical variable Grammar comprised of % Grammatical and % Complex for distinguishing groups (Table 12). The effect of Grammar in relation to Group, controlling for all other variables ($\eta^2_p = .264$), was of large magnitude (Cohen, 1988). The proportion of variance attributable to the association between Group and the canonical variable Grammar, controlling for all other variables, is presented in Figure 6. The proportion of variance associated with Grammar was second only to Informativeness suggesting that it is an important discriminative variable in characterizing the discourse of participants with PD. However, because of the interaction effect separate MANOVAs for the effect of group were conducted for each Task (Picture and Retell) separately to fully examine the effect of Grammar on discriminating between groups. Box's M was not significant for either of the follow-up analyses: Picture ($p = .001$) or Retell ($p = .002$). The effect of group for the Picture task was significant $F(2, 35) = 11.79, p = < .000, \eta^2_p = .402$. However, the effect of group for the Retell task $F(2, 35) = 1.23, p = .304$ was not significant. Specifically, for the Picture task, there was a significant group effect for % Grammatical $F(1, 36) = 20.83, p = < .000, \eta^2_p = .367$ but not for % Complex $F(1, 36) = 3.25, p = .100$. For the Retell task neither the effect for % Grammatical $F(1, 36) = 2.37, p = .132$ nor % Complex $F(1, 36) = .300, p = .587$ were significant. Collectively, the findings from the individual MANOVAs suggest that only % Grammatical differs between groups and that the source of this effect are the group differences observed for the Picture vs. the Retell task.

Common pragmatic errors in the PD included omission of markers for verb tense as in the following example: *The farmer watch him disappear in the distance*. Omission of articles was also frequently observed in the PD group and rarely in the control group: *He stops c-car gets the map out*.

Lexical in relation to Group. The effect of the canonical variable Lexical, comprised of % Open and % Verb for distinguishing between the PD and control groups, was not significant (Table 12). These results suggest that participants with PD did not differ from controls along the dimensions of lexical use as defined by the proportion of open class words and verbs produced.

Informativeness in relation to Group. There was a significant effect of the canonical variable Informativeness, comprised of Total CIUs, % CIUs, and CIUs/Minute, for distinguishing groups (Table 12). The effect size for Informativeness in relation to Group, controlling for all other variables, was large ($\eta^2_p = .451$) (Cohen, 1988). As the results presented in Figure 6 show, the canonical variable Informativeness was associated with the largest proportion of shared variance with Group when compared to the effects of Productivity, Grammar, and Verbal Fragmentation. All of the univariate ANOVAs performed subsequent to the multivariate test for Informativeness were significant indicating that Total CIUs, % CIUs, and CIUs/Minute differed significantly between PD vs. controls (Table 12). The partial eta squared values were interpreted to indicate the presence of large effects of Group for Total CIUs and CIUs/Minute and a medium to large sized effect for % CIUs (Cohen, 1988). As the descriptive data for Group in Table 12 shows, participants with PD produced discourse samples with significantly fewer overall correct information-conveying words (i.e., CIUs) and a lower proportion of information-conveying words (i.e., % CIUs) vs. controls. They also conveyed task-relevant content less efficiently vs. controls (i.e., CIUs/Minute).

Verbal Fragmentation in relation to Group. There was a significant effect of the canonical variable Verbal Fragmentation comprised of Total Disruptions and Disrupt/100 Words for distinguishing PD vs. control groups (Table 12). The effect of Verbal Fragmentation in relation to Group, controlling for all other variables, ($\eta^2_p = .187$)

was interpreted as a medium to large effect (Cohen, 1988). The proportion of variance accounted for by the association between Verbal Fragmentation and Group is presented in Figure 6. In the subsequent univariate tests only Disrupt/100 Words demonstrated a significant effect of Group (Table 12) that is interpreted as medium to large in magnitude (Cohen, 1988). As the descriptive group data reported in Table 12 show, participants with PD produced on average significantly more verbal disruptions per 100 words than controls. The increase prevalence of verbal disruptions in the PD groups was often manifested as complex episodes of disruption. In the following example, the individual with PD experiences a complex disruption that contains pauses, interjections, sound repetitions, and revisions/reformulations in addition to a grammar error:

<s –s –s> [/] see um [pause] driving down the road in a convertible.

And he <stops by> [/] stops by a farmer <w – whose got> [/] whose a <digging a tree> [/] ugh plant a tree

The results of the five MANOVAs and subsequent univariate tests facilitated the development of a set of variables characterizing the discourse differences between participants with PD and controls. These findings indicate that the discourse of participants with PD can be discriminated from that of controls along a continuum of discourse domains and measures including: (a) productivity - specifically total words produced and spoken WPM; (b) grammar - specifically the proportion of grammatically correct and well-formed utterances; (c) informativeness - specifically total CIUs, proportion of CIUs per total words, and number of CIUs spoken per minute; and (d) verbal fragmentation specifically the number of verbal disruption behaviors produced per 100 words spoken.

4.3 Discourse Discriminant Function (RQ 4)

A discriminant function analysis (i.e., DFA) was conducted to determine whether a single canonical discourse variable could, with sufficient sensitivity and specificity, differentiate the spontaneous language of PD vs. controls. The explicit goal, and the primary objective of the current study, was to develop a single canonical discourse function that could

predict whether or not a particular discourse sample belonged to a participant with PD vs. a control. While the multivariate and univariate analyses of variance reduced the discourse variables to a minimal set of measures that differed between groups, the role of DFA was to determine the ideal canonical weighting of those variables (i.e., the discriminant function) for optimizing the separation of groups. In other words, the DFA helped create a true discriminative profile that characterizes the discourse of PD from that of controls.

Developing the model for the discriminant function. For the initial DFA, Group (PD vs. control) was entered as the DV. For the IVs predictor variables, at least one variable was included from each of the significant canonical variables (e.g., Productivity, Grammar, etc.). To determine which variables were entered as predictors, first the MANOVA multivariate results were used to determine which discourse domains (i.e., canonical variables) were able to differentiate between groups. From those domains (i.e., Productivity, Grammar, Informativeness, Verbal Fragmentation), the subsequent univariate tests of between group effects were used to determine specifically which individual variables differed significantly between PD vs. control groups. The IVs, predictors, in the DFA included: (a) Total Words and WPM (Productivity); (b) % Grammatical (Grammar); (c) CIUs, % CIUs, and CIUs/Minute (Informativeness); and (d) Disrupt/100 Words (Verbal Fragmentation).

Testing the discriminant function. Of no great surprise given the earlier MANOVA results, the tests of equality of means were significant ($\alpha = .05$) for all of the IVs confirming their appropriateness for inclusion in the discriminant function (Total Words ($F(1, 74) = 5.30, p = .024$); WPM ($F(1, 74) = 13.07, p = .001$); % Grammatical ($F(1, 74) = 14.35, p < .001$); CIUs ($F(1, 74) = 16.31, p < .001$); % CIUs ($F(1, 74) = 12.50, p = .001$); CIUs/Minute ($F(1, 74) = 22.32, p < .001$); Disrupt/100 Words ($F(1, 74) = 6.35, p = .014$). The correlation between variables was reviewed using the Pooled Within-Groups Matrices output from IBM SPSS Version 22.0 to assess the degree of independent information contributed by each variable within the fitted model, CIUs and Total Words were highly correlated $r = .918$ which is an indicator of singularity between these two variables and thus an assumption violation for DFA. No other pairwise comparisons of

variables were of concern for redundancy with the remaining correlation coefficient values $< .80$. From these results it was concluded that Total Words and CIUs were highly redundant with each other but not with any other variables. For the overall model, Box's M was significant with $F = 4.43, p = < .000$. However, given the n sizes in the current study this was not of significant concern for the integrity of the analysis. Despite the significant Box's M, the log determinants were of similar enough values (PD = 36.49, Control = 29.98, Pooled = 35.10) to mitigate significant concerns over violations of homoscedasticity. The canonical correlation value was .624 indicating that the model for the discriminant function explained 38.94% of the variance between groups. The discriminant function for discourse showed a significant association between groups and predictors ($\lambda = .610, \chi^2(7) = 34.81, p < .001$) with 61.0% of between group variance unexplained after accounting for the proportion of variance attributed to the discriminant function. The overall classification for the fitted model of discourse variables showed that it correctly classified 84.20% of cases into either PD or control groups. The cross-validation classification results also were strong with 78.90% correctly classified cross-validated cases. Despite the strength of the original discriminant function, there were two redundant variables identified (i.e., Total Words and Total CIUs). Redundant variables are of particular concern in DFA because they reduce the stability of the model. Consequently steps were taken to refine the model and remove the redundant variables.

Refining the discriminant function and removing redundancy. In Step 2, two additional DFAs were conducted to determine the optimized model for separating the groups by eliminating the redundant variable contributing the least amount of unique information to the model. Box's M was significant for both DFAs but again was not judged to be of significant concern. The resulting combination of discourse variables for both DFAs were statistically significant indicating that there were significant relationships among the predictors and Group such that both functions were acceptable for discriminating the groups. To facilitate comparisons across models, Table 13 presents a summary of the variance, predictive accuracies, and statistics for the original model and the two test models.

Table 13.

Summary of the Assessed Discriminant Functions

	Original Model	Test Model 1: Total Words Removed	Test Model 2: CIUs Removed
Proportion of group variance accounted for by the model (R^2)	38.94%	38.94%	38.56%
% of overall cases correctly classified (i.e., 'hit ratio')	84.20%	85.50%	81.60%
% of cross-validated cases correctly classified	78.90%	78.90%	77.60%
% of PD cases correctly classified	81.60%	81.60%	78.90%
% of Control cases correctly classified	86.80%	86.80%	84.20%
Statistical test results for Wilks' Lambda Test	$\Lambda = .610$, $(\chi^2(7) = 34.81$, $p = < .001$), $\eta^2 = .390$	$\Lambda = .610$, $(\chi^2(6) = 35.04$, $p = < .001$), $\eta^2 = .390$	$\Lambda = .614$, $(\chi^2(6) = 34.62$, $p = < .001$), $\eta^2 = .386$
Box's M	$F = 4.43$, $p < .001$	$F = 4.8$, $p < .001$	$F = 4.56$, $p < .001$

The results of the DFAs for the two test variables suggested that the removal of the redundant variable Total Words increased the model stability without negatively affecting the predictive properties of the discriminant function. Test Model 1 was selected as the final model for the discriminant function and was renamed PD Discourse. Despite a significant Box's M test for the discriminant function PD Discourse (Table 13), the log discriminant values were acceptable and mitigated significant concerns over violations of homoscedasticity (PD = 31.12, Control = 25.73, Pooled = 29.92). The discriminant function for PD Discourse included the following variables: WPM, % Grammatical, CIUs, % CIUs, CIUs/Minute, and Disrupt/100 Words. The DFA, in keeping with the

between group multivariate MANOVA results, reinforces the importance of informativeness, grammar, and verbal fragmentation in the characterization of PD discourse. These variables carried the highest weightings in the discriminant function. The correlation analysis showed that each of the informativeness variables contributed unique information to the discriminant function with % CIUs and CIUs/Minute carrying more discriminative value in the model vs. CIUs. Measures of productivity carried the lowest weightings in the discriminant function suggesting they are of lesser importance in the profile of discourse impairment in PD without dementia. The discriminant function for separating the spontaneous language of PD vs. controls follows:

- Discourse Discriminant Function =
 $(- .004 \times \text{WPM}) + (.021 \times \% \text{ CIUs}) + (.023 \times \text{CIUs/Minute}) + (.055 \times \text{Disrupt/100 Words}) + (.044 \times \% \text{ Grammatical}) + (.014 \times \text{CIUs}) + (- 9.463)$

With a sensitivity of 81.60% and a specificity of 86.80% the discriminant function PD Discourse was judged to be effective at separating the spontaneous language samples of PD vs. controls. The ‘hit ratio’ exceeded chance by more than 25% with 85.5% of discourse samples correctly classified into PD vs. control groups. A summary of the classification accuracy of the final discriminant function is presented in Table 14.

Summary of Overall and Cross-Validated Predictive Accuracies for PD Discourse

			Predicted Group Membership		Total
			PD	Control	
PD Discourse	Count	PD	31	7	38
		Control	5	33	38
	%	PD	81.60	18.40	100.00
		Control	13.20	86.80	100.00

PD Discourse (cross-validated)	Count	<i>PD</i>	29	9	38
		<i>Control</i>	7	31	38
	%	<i>PD</i>	76.30	23.70	100.00
		<i>Control</i>	18.40	81.60	100.00

The group centroid values were $-.788$ for PD and $.788$ for controls. These scores facilitate the classification of new discourse samples into either PD vs. control groups based on where the Discourse Discriminant Function score calculated for a given sample falls relative to the centroid values. Scores closer to the centroid value for PD (i.e., negative values on the y-axis Figure 7) are classified as PD and vice versa for controls (i.e. positive values on the y-axis). Figure 7 illustrates the effectiveness of the separation of discourse samples for the current study using the discriminant function. As Figure 7 illustrates more classification errors were made for the Retell ($n = 8$) tasks than for Picture tasks ($n = 4$).

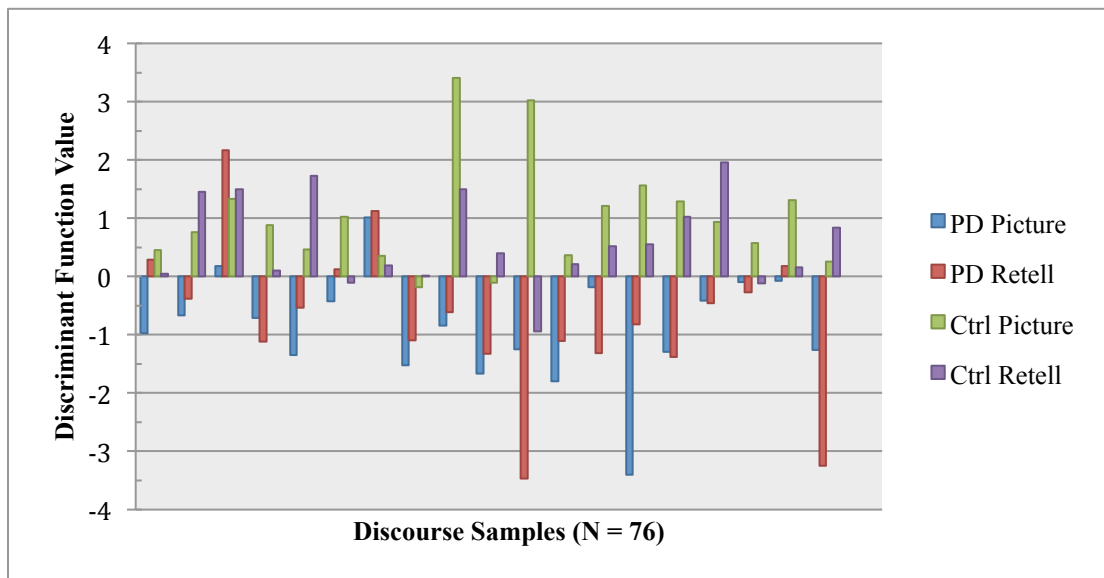


Figure 7. Figure 7 illustrates the classification of individual discourse samples using the Discourse Discriminant Function.

These results confirm that a single discriminant function, comprised of a set of discourse variables, derived from a multi-genre, multi-level discourse analyses approach, can with acceptable sensitivity and specificity correctly classify spontaneous language samples as PD vs. control independent of the discourse task used to elicit the sample. Moreover, the results of the DFA validate a characteristic profile of discourse challenges occurring in the presence of PD in the absence of dementia.

4.4 Relationship Between Discourse Performance and Disease Variables, Age, Education and Cognition (RQ 5)

Using the discriminant function score generated for each discourse sample in the PD group, a series of two-tailed Pearson's Product Moment correlations were conducted to determine the nature and presence of relationship between discourse performance and markers of disease severity specific to PD (i.e., LED, Hoehn and Yahr, UPDRS II, and Duration) and variables reported to affect language performance (i.e., age, education, global cognitive function). The analyses were conducted separately for the Picture and the Retell data to help understand whether the nature of these relationships differed by discourse task.

Relationships among discourse performance and disease severity. For the Picture task, significant correlations ($\alpha = .01$) were observed for LED ($r = -.702$) and Hoehn and Yahr ($r = -.679$). The relationship between the discourse discriminant function score and UPDRS III was significant at a level of .05 with $r = -.532$. The relationship between the discourse discriminant function score and duration of disease was not significant ($r = .070$). These results suggest that for the picture task for participants with PD there was a significant negative relationship between select markers of disease severity and discourse function. This finding means that as disease severity increases (i.e., higher medication needs, increased disease staging, and higher motor severity scores) discourse performance worsens.

For the Retell task, no significant correlations ($\alpha = .01$ or .05) were observed between the discourse discriminant function scores and markers of disease severity for LED ($r = -.009$), Hoehn and Yahr ($r = -.416$), UPDRS III ($r = -.301$), and Duration ($r = .249$). While the

correlations were in the expected direction, they failed to reach significance suggesting that the relationship between disease severity and discourse performance is potentially stronger for the Picture task than for the Retell task.

Relationships between discourse and language performance predictors. A second series of Pearson's Product Moment correlations were conducted using the discourse discriminant scores for all participants, $n = 38$ for PD and $n = 38$ for controls. Separate two-tailed bivariate correlations were conducted for the Picture and the Retell tasks to understand the nature and the degree of relationship between discourse performance and age, education, and global cognition (MDRS-2 scores). For the Picture task, there was a significant positive relationship between global cognition and discourse performance ($\alpha = .01$) such that higher MDRS-2 scores correlated with better discourse performance ($r = .540$). There was no significant relationship observed between discourse performance and age ($r = -.201$) or years of formal education ($r = .128$).

For the Retell task, there was a significant positive relationship ($\alpha = .01$) between global cognition and discourse performance such that higher MDRS-2 scores correlated with better discourse performance ($r = .461$). With $\alpha = .05$ there was a positive significant relationship observed between years of education and discourse performance such that increased years of education was associated with better discourse performance for the Retell task ($r = .359$). There was no significant relationship observed between discourse performance and age ($r = -.056$). Collectively, these results suggest that discourse performance is associated with global cognition for both PD and controls and for both discourse tasks. No relationship between age and discourse performance was observed for either task. Furthermore the results suggest that discourse performances for Retell tasks are more susceptible to differences in years of education whereas, performances on Picture tasks do not share an association with education.

Chapter 5: Discussion

The discussion follows the format of previous chapters. Section 5.1 addresses design and methodological issues. This section is intended to provide a comprehensive discussion of the methodological and participant description considerations relative to the unique contributions of the current study to the body of discourse literature in PD. The subsequent sections present a discussion of the findings from the individual research questions in the context of the existing literature. Sections are then presented on the limitations of the study, clinical implications of the findings and on future directions/considerations.

In the current study, it was hypothesized that changes in cognition and language, occurring in the context of Parkinson disease, affect spontaneous language in a predictable manner that can be uniquely characterized using spoken monologic discourse tasks. The results of this study confirm this hypothesis and advance our understanding of spontaneous language changes in PD in the absence of dementia by developing a profile of discourse impairment in PD that discriminates effectively the discourse of PD vs. controls with greater than 80% sensitivity and specificity. The findings from the study make a novel contribution to the field of communication disorders by expanding our understanding of communication challenges in Parkinson disease beyond the usual spheres of motor speech production and voice that have predominated the clinical view of PD for decades. More importantly, findings from this comprehensive study support a potential methodological paradigm and an analysis tool that can be explored and refined further for applications in PD research and clinical care.

5.1 Participant Sampling and Description Considerations

In the existing PD literature, comprehensive participant descriptions are often lacking which hampers cross-study comparisons and the ability to elucidate the nature of the relationship between discourse performance and relevant disease-specific variables. However, the careful description of participants on variables relevant to performances on cognitive-linguistic tasks, in general, and to PD, specifically, has become even more

important within the recent advances of research examining cognition and language impairments in PD.

Sample size. Sample sizes reported in the existing PD discourse literature have been modest. The mean PD sample size among published studies of spoken monologic discourse is 12.1 with a range of 10 to 21 participants. For controls the mean sample size is 14.9 with a range of 8 to 32 participants. Spoken discourse studies, especially those employing cross-genre, multi-level analyses, require substantial resources placing practical limitations on N sizes. Smaller N sizes, combined with single discourse task sampling practices, create challenges in detecting discourse differences characterized by greater within group variability and smaller effect sizes. Moreover, the dynamic nature of the questions posed in spoken discourse studies requires researchers to analyze samples along a continuum of behaviours creating statistical issues of multiple comparison bias further limiting the detection of smaller, yet potentially important, effects. The current sample size of 38 ($n = 19/\text{group}$) is larger than the mean reported among existing studies of discourse in PD. In the current study, a multivariate approach and an alpha level of .10 were chosen to optimize $1-\beta$ given the inherent resource demands of participant recruitment, data collection, transcription, and analysis. As such, the ability to detect smaller sized effects, possibly obscured in other studies, may partially account for differences in observations between the current study and the existing literature.

Age and education. With few exceptions (Illes, 1989; McNamara et al., 1992), the age and education profiles of participants in the current study were comparable to other studies of monologic spoken discourse in PD (Ash et al., 2011; Ash et al., 2012a; Murray, 2000; Murray & Lenz, 2001). There were no significant differences in the age or education distributions between the PD and control groups. The existing literature in healthy older adults paints a complicated picture of age effects on spoken discourse performances. In healthy adults, Capilouto, Wright, and Wagovich (2005) reported no effect of age for words per minute or CIUs per minute but did report significant differences in % CIUs between younger adults ($M = 22.4(2.2)$ years) and older adults ($M = 71.4(8.2)$ years). Similarly, Mackenzie (2000), using a picture description task in

healthy adults, reported that age effects were significant only for efficiency of communication content (i.e., using a measure similar to % CIUs) among a coterie of microlinguistic and macrostructural discourse variables (Mackenzie, 2000). Moreover, while participants in the 40 to 59 years group were more efficient in expressing content vs. those in the 75 to 88 years group, no effects of age were observed in the 40 to 59 years group vs. 60 to 74 years group. Other studies have reported similar findings suggesting that age has minimal effect on most microlinguistic and macrostructural variables for highly structured discourse tasks except in the old elderly (i.e., > 75 years of age) (Cooper, 1990; Glosser & Deser, 1992; Mackenzie, 2000; Marini et al., 2005; Obler et al., 1994; Wright et al., 2014). The two discourse tasks selected for the current study (Picture and Retell) were chosen because they minimize such age-related effects on discourse performance differences which are more apparent on discourse tasks such as single picture descriptions and personal narratives that encourage tangential language production (Marini et al., 2005). In the current study, it is unlikely that age differences contributed significantly to observed performance differences between PD vs. controls. Nonetheless, careful consideration was given to equating the groups to minimize age effects.

The range for years of education, while typical for the population from which participants were drawn, was fairly large in the current study. Mackenzie (2000) reported that education has a significant effect on spoken discourse performance for picture description tasks in healthy adults for a variety of microlinguistic and macrostructural measures (Mackenzie, 2000). Mackenzie's study (completed in the UK) showed that the effects of education on discourse performance were most pronounced for those individuals not completing a minimal level of secondary school with no post-school professional/specialized training. This was not the case for any of the participants in the current study. As such, education was not expected to have a substantial effect on the interpretation of the data from the current study in the context of the existing PD discourse literature. However, it is interesting to note that in the current study a significant relationship between discourse performance and education was found, but only for the Retell task a finding that will be discussed in greater detail later in the discussion.

Motor and disease severity. Although age and education data are ubiquitous across PD discourse studies, few studies comprehensively describe their participants relative to PD-specific variables characterizing motor and disease severity (e.g., UPDRS-III scores, disease duration). While the degree of and nature of the relationships between disease variables and task performance differs across specific cognitive and language tasks, in recent years multiple investigators have underscored the importance of collecting and reporting core disease variables in the characterization of participants with PD in studies of cognition (Green et al., 2002). Moreover, it is crucial to consider how variables such as disease duration or motor severity interact with discourse performance. For example, performances on verbal fluency tasks among persons with PD decline with increased disease staging (Araujo et al., 2011; Koerts et al., 2012; Obeso, Casabona, Bringas, Alvarez, & Jahanshahi, 2012; Pereira et al., 2009). However, the relationship between verbal fluency performance and motor impairment severity is ambiguous (Pereira et al., 2009; Signorini & Volpato, 2006). The findings of these studies indicate that a single marker characterizing motor/disease severity in PD for the purposes of elucidating patterns of discourse performance is likely not sufficient.

With the noted exception of Ash et al. (2011) and Ash et al. (2012a), the comprehensive reporting of relevant disease variables is a void in the current literature on discourse performance in PD. While 80% of studies reported some form of PD disease staging (i.e., Hoehn and Yahr or Webster Disability Scale values), only 50% reported disease duration and 40% reported motor severity using the UPDRS-III. Moreover, few authors with the noted exception of the Ash et al. studies (i.e., 20% of studies), reported medication profiles (e.g., levodopa equivalents) for participants. Collectively, the existing literature of discourse impairments in PD and the growing body of neuropsychological literature indicate that in studies of discourse the collection and analysis of these PD-specific disease characterizing variables are of equal importance to other more typically collected variables such as handedness, languages spoken, etc.

While, the relationship is less definite than that for global motor impairment and disease duration, there is emerging evidence that side of symptom presentation (i.e., right vs. left) and type of motor symptoms (i.e., akinetic rigid vs. tremor dominant) may play an important role in the profiling of cognitive disruptions (Burn et al., 2006) and specifically discourse performance in PD (Holtgraves et al., 2010) relative to PD subtypes. In the current study, participants were not classified according to side of symptom presentation or symptom subtype for two reasons: 1) the exact constructs for labeling these subtypes remains debated in the literature and 2) the sample size was not sufficiently powered for subgroup analyses. Regarding the first point, much work is yet to be done to validate a reliable construct for defining subtypes in PD which is complicated by several factors such as: 1) the finding that right vs. left symptom presentation is not always clear in individuals with PD who can present with more impairment in the lower extremities on one side of the body but an opposite pattern of involvement in upper extremity motor impairments and 2) individuals with PD can initially present with tremor dominant symptoms which evolve into a more akinetic rigid form of the disease. Recently, the work of Stebbins et al. (2013) formalized identification of PD subtypes based on UPDRS scores. As these subtype constructs are more clearly defined, future studies may wish to capture this important data for consideration in analyses.

The mean UPDRS-III score for the current study (30.26) differed from the means reported in other studies by a range of 6.36 points (Ash et al., 2012b) to 16.59 points (Jaywant & Pell, 2010). Using a previously reported clinically important difference score (Shulman et al., 2010), the participants in the current study demonstrated more severe motor symptoms on a magnitude that was of moderate to large clinical importance (Large CID ≥ 10.8 points) in relation to other discourse studies reporting UPDRS-III scores. However, it is unclear as to whether the differences observed in the UPDRS-III scores among studies are the result of true differences in participant populations vs. systematic error associated with the measuring tool and/or raters. One of the challenges in comparing the UPDRS-III data among studies is its reliability. While intra-rater reliability is reported as high with ICCs $> .90$ (Post, Merkus, de Bie, de Haan, &

Speelman, 2005; Siderowf et al., 2002), interrater reliability is less robust (Post et al., 2005). Comparative studies demonstrate a large degree of variation in UPDRS-III scores such that two different raters assessing the same individual over two separate research visits report score differences of up to 16 points (Post et al., 2005). Post, Merkus, de Bie, de Haan, and Speelman (2005) reported that level of training (i.e., movement disorders neurologist, nurse, and neurology resident) has a significant effect on interrater reliability for the UPDRS with movement disorders neurologists consistently assigning lower UPDRS-III scores vs. nurses and residents. Moreover, UPDRS-III scores can vary substantially between 'on' and 'off' states relative to medicine administration. In the current study, steps were taken to ensure that scores were reliable and optimally interpretable relative to the discourse findings including the use of a single rater, completion of the Movement Disorders Society-Unified Parkinson Disease Rating Scale (MDS-UPDRS) certification exam by the researcher (AR), assessing participants in relatively similar conditions relative to medication timing, and administering the UPDRS-III (i.e., motor assessment subtest of the MDS-UPDRS) on the same day as the discourse data were collected.

Similarly the Hoehn and Yahr scores reported in the PD discourse literature also are quite variable. With a mean score of 2.45, a median score of 2.5 and a range of scores from 2 to 4 on the Hoehn and Yahr scale, the PD participants in the current study are generally comparable to the larger body of PD discourse literature. The majority of studies reported average scores between 2 and 3 out of a possible 5 on the Hoehn and Yahr scale. Ash et al. (2011; 2012a) reported a mean Hoehn and Yahr score of 2.2 ($SD = .7$) for both studies. In contrast, Murray (2000) reported a mean Hoehn and Yahr score of 3 with scores ranging from 2 to 4. Jaywant and Pell (2010) enrolled participants with the least advanced disease with a mean Hoehn and Yahr score of 1.75. Overall, the Hoehn and Yahr scores reported in the existing literature suggest that typical participants in discourse studies in PD have bilateral disease of mild to moderate severity with varying levels of balance difficulties but remain physically independent. Hoehn and Yahr scores tend to correlate with scores of motor severity on the UPDRS (Goetz et al., 2004). As such, collectively the data from the UPDRS-III and Hoehn and Yahr scores indicate that the participants in the current study had more severe disease than the participants

reported in the Ash et al. (2011; 2012a) and Jaywant and Pell (2010) studies, but potentially less severe disease than those reported in Murray (2000).

In the PD discourse literature, disease duration is quite variable. The mean years with disease in the current study (9.34) was higher than for the participants in Ash et al. (2011), Ash et al. (2012a), Jaywant and Pell (2010), Murray (2000), and Murray and Lenz (2001); but, was comparable in years to Huber and Darling (2011), Illes et al. (1988), Illes (1989), and McNamara et al. (1992). Capturing data relative to ‘true’ disease duration is challenging in PD given that there is a considerable (and variable) pre-symptomatic phase of the disease when the pathology of PD is advancing but physical and autonomic symptoms remain at a sub-threshold level impeding the assignment of a diagnostic label using current criteria (Braak et al., 2003; Postuma, Gagnon, & Montplaisir, 2010). Research efforts continue to search for reliable and accurate tools for diagnosing PD in its pre-symptomatic state (Postuma et al., 2010). However, until such time as these tools exist, understanding fully the the relationship between disease duration and discourse performance will remain challenging. Nonetheless, describing PD participants relative to the duration of their PD symptoms, in the context of other disease-specific measures, is informative.

The value of reporting levodopa equivalent dose (LED) in studies of PD is that it provides a marker of the amount of dopaminergic medication required to manage PD symptoms. While not universal, higher LEDs are associated generally with more severe motor symptoms and longer disease duration. Furthermore, levodopa can have varying effects on cognitive and language tasks (Poletti & Bonuccelli, 2013). Although to date, no study has explored ‘on’ vs. ‘off’ effects in spontaneous language performance in PD. As such, LEDs provide a metric for comparing participants across studies and are frequently reported in PD studies of cognition. Only three prior studies of discourse in PD reported LEDs (Ash et al., 2011; Ash et al., 2012a; Ash et al., 2012b). The mean LED for the current study was 984.37 mg with wide 95% CIs suggesting participants were variable relative to dopaminergic medications required to manage symptoms. The mean LED for the current study is substantially higher than the dosing reported in other studies of PD discourse. The difference in LEDs is consistent with the higher UPDRS-III

and higher Hoehn and Yahr scores reported for participants in the current study and may further indicate that they had more severe motor symptoms vs. those reported in the Ash et al. (2011), Ash et al. (2012a), and Ash et al. (2012b). Participant differences in motor severity, reflecting more advanced pathological states of PD, may aid in the interpretation of discrepancies between the current study and the Ash et al. studies that are expounded in the discussion of the results associated with the specific research questions.

Depression. Depression is common in PD occurring in an estimated 30 to 40% of individuals (Schrage et al., 2007). Depression in PD is linked to impairments in cognitive processes, specifically executive functions and confrontation naming for nouns (Poletti & Bonuccelli, 2013; Tröster, Stalp, Paolo, Fields, & Koller, 1995). The presence of depression in the control cohort was screened via self-reported history. In the PD cohort, the depression screening protocol included self-report, a clinical chart review, and item 3 from the UPDRS-I. It was not reasonable, within the limitations and purposes of the study, to access the medical records of controls. The UPDRS is a measure validated in PD only and, thus, would not have been appropriate to administer to the control group. In the Schrage et al. (2007) review of depression screening tools in PD Part I of the UPDRS was determined to be sufficient as a “crude screening tool” for depression in PD (p. 1085). Alpert, Rosen, Welkowitz, and Lieberman reported significant correlations between depression scores and measures of productivity and verbal disruption in the spoken discourse of participants with Parkinson disease dementia (PDD). However, only two studies of spoken discourse in PD without dementia used a formal screening protocol for depression. Murray (2000) and Murray and Lenz (2001) used the Hamilton Depression Scale. An identified limitation of the current study, discussed at length in Section 5.4, is the lack of a formal screening instrument for depression in the control group and the use of a limited measure in the PD group. However the expected potential impact of depression/anxiety factors on the study results may be minimized given that the majority of PD participants were men and that women with PD have a significantly higher risk of depression/anxiety vs. men with PD (Solla et al., 2012; Song, Gu, An & Chan, 2014). Moreover, as discussed in detail in section 5.4, previous studies have not found a significant effect of depression on discourse measures similar to those reported in the current study (Murray, 2010).

Global cognition. Researchers demonstrated that global measures of cognition correlate with performances on more discrete language tasks in PD (Bayles, 1990; Colman et al., 2009; Crescentini et al., 2008; Grossman et al., 2000; Henry & Crawford, 2004; Rodriguez-Ferreiro et al., 2010). Within the existing body of spoken discourse literature in PD, standardized measures of global cognition were reported in 63.6% of studies. Another 27.2% of studies included some version of a self-report task for global cognition. Given the potential influence of global cognition on discourse performance and the elevated risk of dementia in the PD population, controlling for cognition and appropriately screening for PDD is important. Recent studies identified which global measures of cognition are most valid in PD wherein the Mattis Dementia Rating Scale-2 (Mattis, 2001) and the PD-Cognitive Rating Scale were reported to be the more valid and reliable measures for identifying PD-MCI and PD dementia (Pagonabarraga et al., 2008). The Mini Mental Status Exam (MMSE), a commonly used screening measure for global cognition, is not recommended for use in PD because of its poor sensitivity for detecting PD-MCI and PDD (Marras et al., 2013). Of the currently published literature on discourse in PD, only 30% of studies (i.e., Murray 2000; Murray and Lentz 2001; Jaywant and Pell, 2010) have used a PD-specific valid and reliable global cognitive measure as the primary screening tool for determining the presence or absence of dementia.

In the current protocol, none of the participants met the criteria for dementia or mild cognitive impairment using the published MDRS-2 norms for both PD and controls, respectively. For the PD group all participants exceeded the PD-specific MDRS-2 cut-off scores for PDD (Llebaria et al., 2008). Moreover, all participants failed to meet the criteria for PD-MCI using the Level I criteria from the recently published MDS Task Force recommendations (Litvan et al., 2012). Cognitive impairment is present in an estimated 30% of newly diagnosed PD cases (Elgh et al., 2009) and the risk of developing PD-MCI and or PDD increases with disease duration and age (Aarsland, Brønnick, & Fladby, 2011). Moreover, discourse performance (i.e., story grammar and productivity measures) correlates strongly with select measures of executive function (i.e., attention, switching, inhibition) in healthy adults (Cannizzaro & Coelho, 2013). Spoken discourse is a sensitive task for early cognitive changes in dementia and other

neurodegenerative diseases (Ash et al., 2006; Duong, Tardif, & Ska, 2003; Orange & Kertesz, 2000; Roberts & Orange, 2013; Roberts-South et al., 2012; Zraick et al., 2011). In the current study, where the objective was to create a discriminative profile of discourse impairment in PD independent of dementia effects, it was essential that the measures used to identify cognitive impairment be specific to PD and in keeping with current published guidelines. While the mean MDRS-2 scores in the current study are comparable to those reported in Jaywant and Pell (2010), they are higher than those reported in Murray (2000) and Murray and Lenz (2001) by a minimum of seven points for the PD group and two points for the control group. Jurica et al. (2001) suggested that a total score change of two points is a clinically meaningful difference on the MDRS-2 (Jurica et al., 2001). With this in mind, the mean MDRS-2 total score for both PD (140.11) and control participants (141.95) in the current study suggests that participants were less compromised on a global scale of cognition than the participants (i.e., PD and control) in either Murray (2000) or Murray and Lenz (2001).

Summary. The limited reporting of descriptive variables in the PD discourse literature creates additional challenges in comparing the participants in the current study with comparable studies. Recently, researchers studying discourse performance increased their standards of reporting relative to participants with PD, which will aid in future cross-study comparisons (Ash et al., 2011; Ash et al., 2012a; Ash et al., 2012b; Murray, 2000; Murray & Lenz, 2001). In general, the PD participants in the current study presented with more severe motor symptoms, longer disease duration and better (Murray, 2000; Murray & Lenz, 2001) or equivalent (Ash et al., 2011; Ash et al., 2012a; Jaywant & Pell, 2010) global cognitive profiles vs. the participants with PD in other comparable studies. There were no substantial differences in age or education among studies. Based on the literature and the current study protocol, a suggested minimal set of descriptive data would include: age, education, depression screening scores, UPDRS-III scores, Hoehn and Yahr scores, duration of PD, LED, and a scores from a PD-recommended dementia screening tool. However, as raised in this discussion, researchers must optimize the quality of the data reported by increasing their awareness to the limitations of these tools (i.e., the UPDRS-III) and carefully interpreting the meaningful differences among participant groups (e.g., minimal clinical difference data, etc.).

5.2 Discourse Sampling Considerations

One of the weaknesses identified in the existing body of literature describing discourse performances among persons with PD is that investigators have, with few exceptions, applied rigorous sampling methods. Previous methodology studies in discourse showed that larger samples are required for stable data relative to discourse measures, such as CIUs (Brookshire & Nicholas, 1994b). The current study sought to improve upon previous studies in the PD discourse literature by collecting two trials each of two different discourse genres for a total of four discourse samples. The total length of all discourse samples combined exceeded 300 words for each participant with an average number of 529 words per participant. These methodological aspects are congruent with the minimum recommended sample length suggested by Brookshire and Nicholas (1994a).

A review of discourse studies, in both healthy adults and disease (e.g., stroke, acquired brain injury), elucidates the importance of discourse sampling methods that facilitate the collection of robust samples for analyses. Previous studies have suggested that for monologic discourse tasks the size of the discourse sample affects both productivity measures such as words per minute and informativeness measures such as % CIUs with larger samples providing more stable and reliable data (Brookshire & Nicholas, 1994b). Brookshire and Nicholas's previous work in healthy older adults and adults with aphasia indicates that the greatest benefit to cost ratio (i.e., test-retest stability) in monologic discourse sampling is achieved using 4 to 5 discourse samples, totaling 300 to 400 words in length. With the exception of Murray (2000), who used two picture description tasks to elicit her discourse samples, previous discourse studies in PD have rarely employed sampling methods to the rigor suggested by Brookshire and Nicholas. Investigators in the PD literature have typically sampled only single discourse tasks. Such minimalist sampling methods affect not only the reliability of the data generated but also obscure PD vs. control performance differences other than those with large to very large effect sizes (Brookshire & Nicholas, 1994b). Subtle, yet important, discourse performance impairments may not be detected in single task studies resulting in null effects. The

sampling approaches used to date in the literature are one source of the variability seen across studies.

Discourse scholars have suggested that different discourse genres (e.g., picture description, personal narratives) as well as different stimuli (e.g., stimuli complexity across different picture description tasks) present subtle differences in intrinsic cognitive (e.g., memory, sequencing) and/or linguistic (e.g., semantic constraint, syntax complexity) demands (Brookshire & Nicholas, 1994b; Cherney, 1998; Fergadiotis & Wright, 2011; Fergadiotis et al., 2011; Shadden, 1998a; Marini et al., 2005; Ulatowska, Allard, & Chapman, 1990; Ulatowska, North, & Macaluso-Haynes, 1981). Generally, collecting samples across multiple tasks increases the robustness (i.e., representativeness) of the discourse profile (Brookshire & Nicholas, 1994b; Doyle et al., 1995; Marini et al., 2005; McNeil et al., 2007; Murray & Lenz, 2001). Murray (2000) and Murray and Lenz (2001) published on two different discourse tasks. However, because the findings from these two tasks (i.e., picture description, and topic directed conversation) were analyzed and published in the literature separately, there are challenges interpreting from a clinical lens the potential cross-genre effects in discourse performance within a single sample of individuals with PD.

While sampling across discourse genres is important for capturing representative samples of spontaneous language production, selecting equivalent tasks for multiple-genre sampling paradigms also is an important perspective to consider (Brookshire & Nicholas, 1994b; Doyle et al., 1995; Doyle et al., 2000; Fergadiotis et al., 2011; Marini et al., 2005; McNeil et al., 2001; McNeil et al., 2007). The findings from Fergadiotis and Wright (2011) highlight the variability in microlinguistic measures of discourse that can occur across different discourse stimuli (Fergadiotis & Wright, 2011). It is critical to control stimuli both within and across genres. Previous studies of discourse in PD, even when multiple samples were collected to increase reliability, did not use parallel forms of elicitation stimuli.

Brookshire and Nicholas (1994b), in addition to Nicholas and Brookshire (1993), presented a set of parallel stimuli with sufficient intra-rater reliability to produce stable

discourse sampling. These stimuli included: 2 single picture description tasks, 2 picture sequence tasks, 2 personal narratives, and 2 procedural narratives (Brookshire and Nicholas, 1994; Nicholas and Brookshire, 1993). This set of discourse stimuli has a high degree of ecological validity. Doyle, Goda, and Spencer (1995) reported that the % CIUs calculated from discourse samples elicited with the highly structured Brookshire and Nicholas stimuli accurately predicted the proportion of CIUs produced during a less structured conversation task (Doyle et al., 1995). Collectively, the methodological literature in discourse underscores the importance of using a sufficient number of tasks as well as the importance of balancing the number of stimuli across discourse genres when more than one type of task is used.

Although the most commonly used discourse elicitation method in the existing PD literature is personal narratives (i.e., semi-structured interviews), single picture tasks are reported. While single picture description tasks are useful, the language producer is left to deduce the sequence and organization of events, which in older adults and the old-elderly elucidates the presence of tangentiality, a normal age-related difference in discourse (Marini et al., 2005). In contrast the use of story sequences, which explicitly outline the sequence of events among multiple pictures, reduces tangentiality related to the task and mitigates the effect of age in the data (Marini et al., 2005). Story retelling tasks (i.e., asking a participant to listen to a story and retell it in their own words) also are valuable to the understanding of spoken language abilities in that they tax both verbal memory and sequencing abilities within discourse production. Moreover, retell tasks generate longer discourse samples that aid in capturing a more representative sample of discourse performance especially in the domain of discourse productivity (McNeil et al., 2007). This was a primary consideration in selecting the Retell procedure as one of the two genres reported herein. While Brookshire and Nicholas (1994b) did not propose specific tasks for use in story retelling elicitation, Doyle et al. (2000) validated parallel forms of the story retell protocol using the stimuli from Brookshire and Nicholas's (1997) Discourse Comprehension Test (DCT). Doyle et al.'s story retell procedure demonstrated sufficient test-retest reliability, inter-rater reliability of the parallel forms for use in discourse analyses of productivity (Doyle et al., 2000; Marini et al. (2005)),

informativeness (Doyle et al., 2000; Hula, McNeil, Doyle, Rubinsky, & Fossett, 2003; Marini et al., 2005; McNeil et al., 2001), grammaticality (Doyle et al., 2000; Marini et al., 2005), verbal disruptions (Doyle et al., 2000; Marini et al., 2005), and phoneme production accuracy (Doyle et al., 2000). Given the findings of Godbout and Doyon (Godbout & Doyon, 2000) and Bayles (Bayles, 1990) showing increased impairments on measures of informativeness in PD using discourse tasks with higher memory demands, it is prudent to consider the use of story retelling paradigms validated by Doyle et al. (2000) for discourse sampling in PD in addition to the tasks proposed by Nicholas and Brookshire (Brookshire & Nicholas, 1994b; Brookshire & Nicholas, 1994a).

Summary. Collectively, these studies underscore the importance of using multiple parallel forms of stimuli across discourse genres to elicit discourse samples of sufficient size and representativeness to generate reliable data for informing our understanding of discourse related impairments in PD. The current study is the first study in PD to use a true cross-genre sampling approach. The current approach facilitated the collection of a robust and representative sample for profiling discourse impairments in PD.

5.3 Discussion of Findings in Relation to Specific Research Questions

Following is a discussion of the findings relative to RQ 1 through RQ 4. Each research question is addressed separately in the context of the relevant existing PD spoken monologic discourse literature.

RQ 1: On which measures do participants with PD differ significantly from controls using a comprehensive battery of standardized assessments of cognition, expressive language, and speech intelligibility?

In the current study, participants with PD performed similarly to controls on a variety of standardized measures evaluating frontal lobe functions and language. While in general, the scores of participants with PD were lower than controls on most of the cognitive measures these differences reached statistical significance for only two measures: verb naming: TAWF and action verbal fluency. The findings relative to the measures of frontal lobe functions contrast with a recent meta-analysis whose authors concluded that individuals with PD were more impaired vs. controls on a broad range of frontal lobe

tasks including the Trail Making Test, Stroop, and phonemic verbal fluency (Kudlicka, Clare, & Hindle, 2011). The mean age of the participants in the current study was higher than the value cited in Kudlicka, Clare, and Hindle (2011) ($M = 63.54$ for TMT, $M = 65.02$ for Stroop (interference minus baseline); and $M = 63.99$ years for phonemic fluency). With the vulnerability of frontal lobe tasks to age effects, the older age of the PD participants, and more importantly the controls, potentially reduced the magnitude of effect of group differences on these measures making them harder to detect in the current study. However, an alternative explanation for the conflicting findings is that the meta-analysis conducted by Kudlicka et al. did not specifically exclude studies of individuals with PD-MCI, only studies of PDD (Parkinson disease dementia). As such, their significant findings may include participants with variable cognitive profiles. In the current study, PD participants did not meet criteria for PD-MCI (Level I criteria) (Litvan et al., 2012) or PDD (Llebaria et al., 2008). Consequently, the observed discrepancies in studies may result from differences in overall cognitive profiles of the participants. Using smaller sample sizes than the current study ($n = 12$ PD, $n = 12$ controls), Bohlhalter, Abela, Weniger, and Weder (2009) reported significant group differences between PD vs. controls on a measure of episodic memory similar to the list-learning paradigm used in the ABCD-VL. Similar to the frontal lobe measures, Bohlhalter et al.'s (2009) findings conflict with the results of the current study. However, comparison with the Bohlhalter et al. study faces similar concerns of age and mixed cognitive profiles. The mean age of participants and controls in the Bohlhalter et al. study was substantially younger than in the current protocol ($M = 59.1(7.1)$ for PD; $M = 46.6(10.0)$ for controls). Notwithstanding the value of these studies, the performance of the PD participants in the current study on a coterie of frontal lobe measures is consistent with that reported in other carefully controlled studies of discourse in PD where PD vs. control differences on standardized measures of frontal lobe function failed to reach significance (Ash et al., 2011; Ash et al., 2012a).

Notably, while the assessment battery included a measure of verbal episodic memory, it did not include a specific measure of working memory. However, Engle (2002) postulated that working memory, as a construct, reflects not capacity in a traditional memory sense per se, but attentional control in terms of inhibiting distractions to

maintain focus/access to a target. Consequently, Engle (2002) argued that Stroop tasks, particularly the interference condition, also reflect working memory abilities. As such, while a traditional working memory task such as a digit span task was not administered, aspects of both attentional control relative to working memory and memory capacity in terms of episodic memory were assessed. Moreover, researchers investigating both healthy adults and a variety of disease conditions suggest that spoken discourse performances for tasks similar to those in the current protocol are highly correlated with episodic memory more so than working memory (Ash et al., 2012a; Light & Anderson, 1985; Murray, 2000; Peach, 2013). The ABCD-VL was included in the protocol to evaluate potential group differences between PD vs. controls relative to episodic memory that could affect discourse performances on the story retelling tasks.

While the presence of cognitive impairment in PD is generally well accepted, the presence of true language-based impairments continues to be debated. The results of the current study suggest that participants with PD performed more poorly vs. controls on measures of verb retrieval. The finding of verb specific deficits in both verb confrontation naming and in action verbal fluency is in keeping with a growing body of literature demonstrating deficits in verb retrieval in PD in the absence of significant cognitive decline (Bertella et al., 2002; Boulenger et al., 2008; Cotelli et al., 2007; Herrera & Cueto, 2012; Peran et al., 2009; Piatt, Fields, Paolo, Koller, & Tröster, 1999a; Rodriguez-Ferreiro et al., 2009). To the researcher's knowledge this is the first study of discourse in PD to report measures of verb retrieval. In previous studies of cognition in persons with PD, performances on both verb confrontation naming and on action fluency tasks have shown strong associations with measures of executive function leading researchers to suggest that the challenge retrieving verbs in PD reflects executive function vs. pure language impairments (Crescentini et al., 2008; Peran et al., 2009; Piatt et al., 1999b; Piatt et al., 1999a). Others hypothesized that the deficits in verb naming observed in PD are related to difficulties accessing semantic information associated with movement aspects of verbs resulting from disrupted motor pathways (Bertella et al., 2002; Boulenger et al., 2008; Herrera et al., 2012; Rodriguez-Ferreiro et al., 2009). In the current study, the presence of large effect sizes for verb confrontation naming and medium sized effects for action fluency, occurring in the presence of non-significant

group differences for other measures of frontal lobe functions, is consistent with the evidence showing that verb retrieval tasks in PD may be sensitive to disruptions of larger networks of brain regions associated with semantic, motor and frontal lobe processes (Pulvermüller, 2005).

Only three other published studies specifically addressed semantic knowledge in PD using The Pyramids and Palm Trees test (Ash et al., 2011; Ash et al., 2012a; Péron et al., 2009). While the corrected alpha level applied in the current study did not show significant differences between PD vs. controls, the statistic for The Pyramids and Palm Trees test met criteria for significance with an uncorrected alpha level of .05. Ash et al (2012a) reported significant group effects for The Pyramids and Palm Trees test in their study of discourse production in PD. However, in a companion study Ash et al. (2011) did not find significant group effects on this measure. Péron et al. (2009) reported similar results in a theory of mind study in PD. The differences and inconsistent findings among the existing studies, including those of the current study, indicate a smaller effect size for this measure that is minimally detectable with the *N* sizes typically employed in PD discourse studies. Despite the variability in findings in the literature, PD performance on measures of semantic knowledge in the context of discourse studies are of importance given that preserved semantic knowledge (i.e., the internal representation of a concept on to which a word must be mapped) is essential for on-line word retrieval and use.

In the current study there were no significant differences observed between PD vs. controls on the BNT and on animal naming (i.e., a measure of semantic fluency). The current findings are in general agreement with other studies showing that impairments in word retrieval for objects/nouns, as measured by the BNT and semantic fluency tasks, are more pronounced in PD in the presence of cognitive decline, depression, and advanced motor symptoms (Cooper, Sagar, Jordan, Harvey, & Sullivan, 1991; Green et al., 2002; Tröster et al., 1995). The rigorous inclusion/exclusion criteria used in the current study screened for the presence of dementia and depression so the lack of group effect on these measures is expected and not surprising.

Additionally, in the current study, there were no PD-specific impairments detected using a standardized measure of syntax production (i.e., SPPT). While multiple researchers reported syntax processing deficits in PD (Grossman et al., 1992b; Grossman, 1999; Hochstadt, 2009; Natsopoulos et al., 1993), few investigators examined syntax production in persons with PD. This is the first PD discourse study to report a standardized measure of productive syntax. The results of the current study suggest that individuals with PD are able to access a large variety of syntax structures during a primed production task such as the SPPT, a finding that is consistent with the lack of group differences observed in the proportion of complex syntax structures from the discourse data. Troche and Altmann (2012) used an experimental task, similar to the SPPT, in which participants with PD produced sentences to describe picture stimuli that were designed to elicit complex syntax structures varying across two levels of complexity (Troche & Altmann, 2012). The authors reported that individuals with PD committed more grammar errors (i.e., errors of correctness and completeness) as a function of increased grammatical complexity. Their findings help explain the pattern of deficits observed in the discourse data whereby complexity of syntax structure is preserved in the presence of PD-specific impairments in grammatical accuracy and completeness, a finding discussed in detail later in this chapter.

Summary. One of the challenges in the current body of discourse literature in PD is that few investigators characterized rigorously the cognitive and language profiles of all participants. The current study used a comprehensive battery of tests of cognition that included core assessments recommended for evaluating the presence of mild cognitive impairment in PD (Litvan et al., 2012). In addition, an expanded battery of language measures informed by the relevant literature in PD and the researcher's clinical experience was used. The language measures used, with the exception of the BNT and verbal fluency, have not been previously reported to a large degree in PD therefore their sensitivity and specificity for identifying language differences between the two groups (PD vs. control) is not established. However, all of the measures used, including the SPPT from the Northwestern Assessment of Verbs and Sentences (Thompson, 2011) have been validated and commonly used in neurodegenerative disorders outside of stroke acquired aphasia and in healthy controls. While the sensitivity and specificity of the

individual measures varies, the goal of the battery was to evaluate any significant differences in cognition and language that could influence discourse performance. Relative to **RQ 1**, participants with PD differed from controls on key measures that reflect integrated cognitive and language processes including: verb confrontation naming and action fluency that would predict disruptions in informativeness and verbal fragmentation at the discourse level. In addition, the standardized testing suggested that the ability to produce complex syntax structures would be preserved. Indeed, the profile that emerged from the current study indicates that discourse impairments in PD are predominately characterized by reduced informativeness and increased verbal fragmentation with the proportion of syntactically complex utterances playing no role in the discrimination of discourse samples between PD and controls. These results indicate, as was hypothesized, that discrete impairments on standardized cognitive-linguistic measures are manifest in the spontaneous language tasks of individuals with PD.

Overall, the cognitive and language profiles of participants in the current study are consistent with the profile of cognitive changes in PD in the absence of PD-MCI/PDD. With few exceptions, the cognitive and language profiles of participants with PD in the current study are consistent with other studies of discourse in PD for which similarly rigorous data are available for comparative purposes (Ash et al., 2011; Ash et al., 2012a; Murray, 2000; Murray & Lenz, 2001). The findings presented here underscore the importance of collecting comprehensive measures of cognition and language in discourse studies in PD for the purposes of comparing results across studies, but more importantly for informing patterns of cognitive-linguistic impairments that manifest in the context of spoken monologic discourse tasks.

RQ2: Does performance on measures of discourse productivity, lexical use, grammar, informativeness, and verbal fragmentation differ significantly between PD vs. controls as a function of discourse task using a cross-genre sampling method and a multi-level discourse analyses paradigm?

While differences were observed between tasks, the lack of interaction effects for 4 of 5 multivariate analyses suggests that, with the exception of grammar, task-specific effects

influenced the discourse of PD and control participants in a consistent way. Relative to **RQ2**, analyses found that performances on measures of discourse did not differ significantly between PD and controls as a function of the discourse task (i.e., Picture vs. Retell) for all of the canonical variables except Grammar.

Analyses of the two discourse tasks (Picture vs. Retell) indicated that the tasks, which differed primarily along a continuum of memory demands (i.e., > for Retell), language model/support (i.e., > for Retell), and visual support (i.e., > for Picture), differentially affected measures of discourse performance. Significant task differences were observed for measures of productivity, lexical use, informativeness/content, and verbal fragmentation. For measures of grammar, controls were affected by Task such that a higher proportion of correct and well-formed utterances were produced on the Picture vs. Retell task. However, the PD group was not affected by task. Two potential explanations emerge for this finding. The first explanation is that individuals with PD experience a ‘maximum hit’ for grammar accuracy/completeness on the Picture task such that their maximum threshold of impairment was reached on the Picture task and further decrement in performance was not experienced with the Retell task as was seen in controls who performed at near ceiling levels for the Picture task relative to % Grammatical. This finding is supported by the significant effect of group on the Picture task but not the Retell task for the canonical variable Grammar and for the individual variable % Grammatical. Secondly, and a theory related to ‘posture first’ theories of motor performance in PD, is that individuals with PD used simpler syntax structures in the Retell task to accommodate for the added task demands which resulted in no change between the two tasks relative to grammar errors or proportion of well-formed utterances; assuming that more paragrammatic errors would be observed on more complex sentence structures as reported by Troche and Altmann (2012). However, this adaptation theory, commonly discussed in the motor literature in PD, is not supported by the data in the current study given that the proportion of complex C-units did not differ significantly between the Picture vs. Retell tasks for either the PD or the control group. While the interaction effect for Grammar is intriguing, especially in light of the role of the basal ganglia in rule-based systems such as grammar, the exact reason for the differences in task effects between the PD vs. control groups remains unresolved. The effect of Task on

grammar accuracy and well-formedness may differ for PD vs. controls such that group differences are more discernable using picture sequence tasks. The findings relative to Grammar in the current study and specifically % Grammatical may aid in disambiguating the existing literature relative to grammar impairments in the spoken discourse of individuals with PD whereby performance differences between PD vs. controls have been reported on some but not all discourse tasks.

Two primary patterns emerged from the analysis of Task effects: a) the Retell tasks elicited discourse samples that were overall more productive relative to volume based measures of discourse (independent of discourse domain) and b) the Picture tasks elicited discourse samples that were more efficient and had a higher proportion of task-specific content words relative to total words spoken. The implication(s) are that examining discourse in PD from the perspective of only a single discourse task/genre affects observations such that productivity effects may be obscured in a picture description task (i.e., or conversely amplified in the Retell task). Likewise, informativeness/content and effects may be obscured using retell procedures. The differential pattern of task effects across discourse domains reflects the differences in communication demands/tasks in everyday language and reinforces the importance of collecting cross-genre data when characterizing or profiling discourse impairments is the primary objective.

The significance of these findings is that the Retell task affected productivity and volume across discourse domains with participants producing a higher number of total words, longer C-units, more total CIUs, and more total verbal disruptions when compared to the Picture task. These findings are consistent with those of McNeil, et al. (2007), who reported that compared with other discourse elicitation methods (e.g., picture sequence description, procedural) story retell procedures yield larger discourse samples consequently affecting measures of productivity. Conversely, the Picture task affected efficiency and content measures with both PD and control participants producing significantly more % CIUs and CIUs per minute for the picture sequence description task than for the Retell task. The prevailing view in the discourse literature is that picture description tasks provide a high degree of semantic and lexical constraint that may result in increased demands for word retrieval (Shadden, 1998a; Shadden et al., 1991).

However, in the current study, both the proportion of and efficiency of production of CIUs (i.e., words accurate and specific to the topic/task) were higher in the Picture task, which may suggest, paradoxically, that the retrieval of specific and accurate words was more challenging in the Retell task. The presence of visual support to aid both memory and word retrieval processes was a fundamental difference between the two discourse tasks which may have acted as a scaffold for word retrieval increasing the proportion of content-accurate and specific words. This explanation is consistent with a hypothesis put forward by Fergadiotis and Wright (2011) in their publication on variations in lexical diversity across discourse tasks (Fergadiotis & Wright, 2011). However, there is a possible alternate explanation to consider. Given that the number of verbal disruptions was significantly higher for the Retell task, an alternate explanation for the reduced efficiency observed in the Retell vs. Picture task is that increased disruptions in the form of revisions, repetitions, and empty words simply may have increased the total number of words without a proportional increase in % CIUs. Another explanation still, and one that is in keeping with the results reported by Illes et al. (1988), is that the picture sequence task may have facilitated more ‘list-like’ descriptions of the events and characters in the picture and less of a narrative (i.e., story-telling) structure, which may have inflated the % CIUs observed in the Picture task. Further detailed analyses are warranted to explore this possibility.

While there was no Task x Group interaction for Verbal Fragmentation, significant task effects emerged with a higher rate of total verbal disruptions on the Retell task vs. the Picture task. Interestingly, no task effects were found for the variable Disrupt/100 words suggesting that the increase in total verbal disruptions is likely a function of the longer discourse samples elicited for the Retell task vs. any cognitive or language effect of task per se. These findings are consistent with the discourse literature in healthy adults wherein researchers reported no effect of task on verbal disruptions when measured as a proportion/100 words (Roberts et al., 2009).

Summary. Collectively, the findings from the task analyses indicate that the relationship between task demands and discourse production are complex. The complexity underscores the importance of understanding cross-genre effects on discourse

and the value of using cross-genre sampling methods to increase the representativeness of discourse samples used in research. As such, the complexity of discourse task effects warrants further examination across the spectrum of Lewy body disorders. These findings indicate that the language sampling practices used in the current study were sufficiently diverse to provide a robust and representative sample of discourse performance for abstracting features that characterized discourse impairment in PD, without the added cost of an interaction effect for most variables.

RQ 3: For which domains (productivity, grammar, lexical, informativeness, and verbal fragmentation), and on which specific discourse measures, do participants with PD differ significantly from controls using a cross-genre sampling method and a multi-level discourse analyses paradigm?

The overarching objective of the current study was to inform knowledge of spoken language in PD by creating a profile of discourse performances in PD. The large variability in sampling practices and analysis methodologies in the existing published literature created obstacles in informing the development of a reliable and testable profile of discourse impairments in persons with PD. The first step in creating a testable discriminant function for discourse in PD was to collect a rigorous corpus of cross-genre, multi-level discourse data that could be analysed to extract a set of discourse variables characterizing the spoken monologic discourse differences between PD vs. controls. These analyses revealed that PD differed from controls along the dimensions of both microlinguistic and macrostructural aspects of discourse performance including productivity, grammar, informativeness, and verbal fragmentation. There was no difference between the discourse of individuals with PD for lexical use variables including the proportion of open class words and the proportion of verbs used. Relative to **RQ 3**, the findings from the current study suggest that the profile of discourse impairment in PD is characterized by: a) reduced productivity marked by fewer total words and fewer words per minute, b) reduced proportion of grammatically correct and well-formed utterances, c) increased density of verbal disruptions per 100 words, and d) reduced informativeness marked by fewer total CIUs, a lower proportion of CIUs per

total words, and fewer CIUs per minute. A discussion of the significant findings and implications of these performance differences follows.

Productivity. Spoken language productivity depends on a variety of variables including, although not inclusively, the intention of a communication event, the outcome of the communication event, the communication environment, the choice of topic, and the listener's perspective, among other influential factors (Searle, 1969). Spontaneous language that is efficient in terms of 'flow' of words in the time domain may be judged as acceptable on the basis of productivity even if the total volume of output is inadequate. Conversely, impaired efficiency of verbal output can have a negative impact on a listener's understanding even if the total output is sufficient. In the current study, the discourse of participants with PD was marked by both impairments in reduced volume of output and impairments in reduced efficiency of output.

Individuals with PD performed more poorly than controls on 2 of 3 measures of productivity: Total Words and Words/Minute. There were no group differences observed for the variable Words/C-unit. The finding of reduced total words is consistent with the results published by McNamara et al. (1992) who reported that when measured in total words, the discourse productivity of participants with PD was more impaired than both healthy controls and participants with Alzheimer's disease (McNamara et al., 1992). However, this contrasts with other studies that have reported no significant effects of PD for total words spoken (Ash et al., 2011; Crucian et al., 2001; Murray, 2000). Aside from total volume of output the efficiency of spoken output can be measured by calculating words per minute (WPM) with discourse sample duration as the denominator. In the current study, total speaking time (including pauses > 2 seconds) was entered as the denominator in the WPM equation. Although studies in PD discourse showed significant differences in WPM between measures that include and those that exclude long pauses (Ash et al., 2012a; Illes et al., 1988), for the current study the researcher was interested in capturing the effect of such pause behaviours on the characterization of discourse impairments in PD. Consistent with the findings of the current study previous authors reported that individuals with PD produce fewer WPM vs. controls (Illes et al., 1988; Illes, 1989). However, other researchers reported that efficiency deficits in PD emerge

only in the presence of dementia (Ash et al., 2011; Ash et al., 2012a). While the approaches differed, both the current study and the Ash et al. studies undertook rigorous processes to ensure that PD participants without dementia were correctly categorized. Notwithstanding the results reported by Ash et al., the finding of impaired efficiency in the current study in the absence of PDD or PD-MCI would suggest that a cognitive explanation for productivity impairments in PD is not sufficient in isolation.

Words/C-unit is a measure of productivity that is influenced highly by grammatical complexity. Longer utterances tend to be more grammatically complex and as such mean length of utterance (i.e., MLU) often has been used as a proxy measure for grammatical complexity. There were no differences in words per C-unit in the current study suggesting that while overall volume of output and efficiency of output were affected, the length of each minimal language unit was not affected in PD. The non-significant difference between groups for words per C-unit in the current study aligns with the finding of a non-significant difference between PD vs. controls on the percentage of grammatically complex utterances. These findings are in keeping with the existing body of literature reporting the absence differences in mean length of utterance between PD vs. controls (Ash et al., 2011; Murray, 2000; Murray & Lenz, 2001). Interestingly, Holtgraves et al. (2010), using personal narratives, reported utterance length differences but only in individuals with PD with left sided motor asymmetry vs. those with right sided motor asymmetry. The Holtgraves et al. finding raises interesting questions relative to the relationship between PD phenotype and spontaneous language production.

Significant findings for both Total Words and WPM, in the context of significant group differences in the number of disruptions/100 Words (expounded upon later in this section), suggest that in isolation or in integration with other discourse processes, non-word verbal disruptions (e.g., pauses exceeding two second duration, non-verbal fillers such as ‘uh’ and ‘uhm’) play a substantial role in productivity impairments associated with PD discourse. Several researchers reported that unfilled pauses between clauses and utterances contribute to the productivity impairments observed in the discourse of individuals with PD (Ash et al., 2011; Ash et al., 2012a; Illes et al., 1988; Illes, 1989). The interaction between non-word verbal disruptions and productivity in the discourse of

PD occurs across a variety of elicitation methods suggesting that this may be a universal feature of discourse impairment in PD.

Productivity measures, more so than other discourse variables (with the exception of verbal disruptions), can reveal motor speech aspects of spoken discourse. The simplest explanation for the observed productivity impairments is a motor explanation given that PD is primarily a motor disease. The assumption is that differences between PD vs. controls for total words and WPM are the result of motor speech impairments in the PD group. However, in the current study there were no baseline speaking rate differences between PD and controls. Moreover, there were no speech intelligibility differences between the two groups. While motor speech deficits could account for both a reduction in volume and efficiency of words produced, it is unlikely given the equivalent findings on motor speech measures that this is the sole source of productivity disruptions in PD. Furthermore, Murray (2000) reported no significant correlations between total words and speech intelligibility further indicating that motor speech differences alone are unlikely to account for productivity impairments in PD discourse. While the Ash et al. studies (2011; 2012a) did not report baseline speech rate or intelligibility measures, they reported no significant differences in UPDRS-III scores between their productivity-unaffected PD group without dementia and the dementia group who did exhibit productivity impairments. The lack of difference in the UPDRS-III scores would suggest that overall motor severity did not account for the productivity differences reported between those two groups. Consequently, while the overall motor severity of PD participants in the current study was greater than the motor severity of participants in the Ash et al. studies it is unlikely that the differences in motor severity accounted for the discrepancy in productivity results among studies.

Given the findings of the methods paper published by McNeil and colleagues (2007) showing that the story retell procedure results in discourse samples with higher volume-based productivity measures vs. other discourse elicitation methods, it is more likely that language sampling differences account for the discrepancies observed in productivity among discourse studies in PD. Specifically, the findings from the task analysis in the current study suggest that the inclusion of a story retell task increased not only the overall

length of the discourse samples in the current study but also their representativeness especially relative to productivity measures. The sampling approach used in the current study revealed group differences in productivity not typically manifested on the single genre, single task sampling methods used in other PD discourse studies. The findings of the current study, contextualized in the existing body of PD discourse literature, support the conclusions made by others that the productivity impairments observed in the discourse of PD likely reflect challenges in the integration among cognitive, language, and motor processes within the context of spontaneous language (Illes et al., 1988; Illes, 1989; McNamara et al., 1992).

Grammar. The interest in productive grammar impairments in the discourse of individuals with PD emerged from the work of multiple researchers reporting that individuals with PD demonstrate impairments in processing complex syntax structures (Grossman, 1999; Grossman et al., 2003; Hochstadt et al., 2006; Natsopoulos et al., 1993). In the current study, Grammar played a significant role in discriminating PD vs. controls accounting for almost 27% of the variance observed and was second only to Informativeness in its effect size. However, further examination of the interaction effects showed that the significant group differences in Grammar resulted solely from the effect of Group on % Grammatical for the Picture task. In the PD discourse literature, a variety of measures assessed grammar accuracy and grammar complexity. Multiple researchers reported mean length of utterance as an indicator of grammatical complexity (Ash et al., 2011; Huber & Darling, 2011; Murray, 2000; Murray & Lenz, 2001). While the validity of MLU as a proxy measure of language complexity is accepted in child language research, its validity in adults for the same purposes is questioned (Scarborough et al., 1991). Moreover length of utterance, or in the case of the current study the length of a C-unit, is influenced by word-based verbal disruptions such as repetitions and revisions that can be salient features of discourse impairment in individuals with neurological injury (Davis & Maclagan, 2010; Jokel, De Nil, & Sharpe, 2007). Words per C-unit were considered in the context of the current study a reflection of productivity as influenced by grammatical complexity vs. a measure of grammatical complexity in isolation.

Researchers also extended the study of grammar in PD spoken discourse to include judgments of the accuracy and form of an utterance specifically whether or not utterances meet accepted syntactic rules for forming sentences. Measures of grammatical complexity are commonly reported in the PD literature and typically include measures of the presence of or number of subordinate clauses attached to the main clause of an utterance (Ash et al., 2011; Ash et al., 2012a; Illes et al., 1988; Illes, 1989; Murray, 2000; Murray & Lenz, 2001). Similarly, the existing PD discourse literature included measures of the proportion of utterances with grammatical errors and/or utterances that are not ‘well-formed’ following the conventions of the language in which the study was conducted (Ash et al., 2011; Ash et al., 2012a; Illes et al., 1988; Illes, 1989; Murray, 2000; Murray & Lenz, 2001). A similar approach was adopted in the current study. Each C-unit was scored using a dichotomous scoring system (i.e., ‘yes’ or ‘no’) for two separate measures of grammar/syntax: a) grammatical complexity scoring the presence or absence one or more adjunctive or dependent clauses in the C-unit and b) grammatically well-formed C-units scoring the presence or absence of grammatical errors using established conventions (see Appendix H).

There were no significant differences in grammatical complexity in spoken discourse between PD vs. controls for either task. The finding that individuals with PD do not produce less complex syntax structures across a variety of spoken monologic discourse tasks is a consistent finding in the literature (Ash et al., 2011; Ash et al., 2012a; Illes et al., 1988; Illes, 1989; Murray, 2000; Murray & Lenz, 2001). Strengthening this finding is the fact that participants with PD did not demonstrate impairments on a standardized measure of syntax production vs. controls. The emerging evidence from the discourse literature suggests that PD-specific deficits in processing complex syntax structures do not correspond to deficits in the proportion of complex structures produced in the spoken language domain. The discrepancies observed in syntax between the comprehension and production domains strengthens the hypothesis asserted by some researchers that syntax disruption in PD does not result from a central process of impaired access to syntax structures but instead reflects disruptions to the underlying cognitive processes (e.g., memory, allocation of attention resources) that support syntax processing (Grossman et al., 1992a; Grossman et al., 2000; Grossman et al., 2000b). In spoken discourse,

researchers reported previously that cognition interacts with syntax complexity. Ash et al. (2011) and Ash et al. (2012a) reported that participants with PDD and those with dementia with Lewy bodies produced less complex utterances than either controls or individuals with PD without dementia. Furthermore, Illes et al. (1988) reported that syntax complexity interacted with disease severity such that participants with more severe disease/disability produced less complex utterances. Collectively, the findings from the current study in the context of the existing body of literature suggest that while the proportion of syntactically complex utterances may not distinguish the discourse of PD vs. controls, it may be an important variable to explore in the context of discriminating and potentially predicting patterns of discourse impairment in PD-MCI and PDD.

While participants did not differ from controls in terms of syntax complexity, they differed significantly in the proportion of grammatically correct C-units that were produced. In other words, individuals with PD produced C-units of equal complexity to controls, but produced more grammar errors than controls. While there is general agreement in the PD discourse literature relative to grammatical complexity, the findings relative to grammatical accuracy and well-formed structures are more ambiguous. Using a picture description task, Murray (2000) reported that individuals with PD produced a lower proportion of grammatically correct and complete utterances compared with the discourse of controls. In contrast, Murray and Lenz did not find PD vs. control differences in grammaticality using a personal narrative task. Furthermore, using a story generation task with picture supports, neither Ash et al. (2011) nor Ash et al. (2012a) found differences between PD vs. controls in the accuracy and completeness of utterances. Although the specific methodologies differed among studies (including the current study), the applied definitions for grammatical accuracy and completeness overlap to a degree such that it is unlikely that the discrepancy in findings result from different construct definitions.

Extending the findings from the grammatical complexity results in the Ash et al. studies, another possible explanation for cross-study differences is that the cognitive status of participants differed across studies. Comparing the cognitive status of the participants in

the current study to those in the Ash et al. studies (2011; 2012a) is challenging because different cognitive measures were used. However, both studies employed stringent dementia screening procedures making it unlikely that there were substantial differences in the cognitive status of participants between studies. This minimizes the likelihood that cognitive differences among the participant groups in the various studies accounted for the conflicting results.

An alternate explanation is that differences in the tasks used among studies may result in differences in results relative to grammatical accuracy and completeness (i.e., picture description, personal narratives, picture sequence descriptions, and story retelling).

Indeed, Murray (2000) found significant differences in grammar accuracy and completeness for discourse samples generated using a picture description task but did not find differences using a personal narrative task (Murray, 2000; Murray & Lenz, 2001).

As discussed in the previous section, the results of the current study suggests that for individuals with PD, grammar performance differences compared to controls may be uniquely revealed using very highly structured discourse tasks such as picture sequences which may, to a degree, disambiguate the findings in the existing literature from Murray (2000) and Murray and Lenz (2001).

Disease severity offers yet another alternate explanation. The participants in the current study exhibited more severe motor symptoms vs. the participant groups in either of the Ash et al. studies (2011; 2012a). In fact, Illes et al. (1988) reported that syntax complexity not only effectively discriminated the more severe PD group from the less severe PD group but also correlated significantly with measures of disease severity (i.e., the Webster 30-point Scale of Parkinsonian Disability) and speech intelligibility.

Similarly, Murray and Lenz (2001) whose participants also presented with more advanced disease vs. the Ash et al. studies, reported significant positive correlations between a measure of speech intelligibility and multiple discourse measures of productive syntax and grammar. Collectively, the findings of the current study and the existing literature in PD discourse indicate that impairments in grammatical accuracy and completeness as observed in spontaneous language production in PD can be associated with motor severity. Given the relationship between overall motor severity and

degradation of the dopaminergic pathways in PD, this raises interesting questions relative to the role of the basal ganglia and dopaminergic system in grammar production in spontaneous language. Teichmann, Darcy, Bachoud-Levi and Dupoux (2009) suggested that the striatum are involved in rule based morphology. Moreover, the left prefrontal cortex which receives inputs from the thalamus that are modulated by GPi and SNr outputs is involved in processing and monitoring aspects of verb agreement and morphology (Kielar, Milman, Bonakdarpour, & Thompson, 2011).

One interesting question that has yet to be addressed in the current corpus of discourse research in PD is whether or not an interaction exists in spontaneous language between complexity of syntax structure and grammar errors. In other words, within spoken discourse tasks, do individuals with PD make more errors on complex grammatical structures vs. simple grammatical structures? There are few experimental tasks of syntax production in PD outside of the domain of spoken discourse. However, a recent study of sentence production using a continuum of complex syntax structures indicates that individuals with PD differ from controls in the number of grammatical errors and in the number of verbal disruptions but only when producing more complex syntax structures (Troche & Altmann, 2012). While in the current study the overall proportion of grammatically complex utterances did not differ between PD vs. controls for either discourse task, the work of Troche and Altmann (2012) underscores the importance of examining the potential interaction between type of syntax structure used and location of grammatical errors in spoken discourse in PD. Such explorations could help inform therapeutic interventions such as counselling individuals with PD to use simpler syntax structures to improve overall grammatical accuracy and completeness in spoken language.

Lexical use. There were no significant differences in either the proportion of open class words or the proportion of verbs in the discourse of individuals with PD vs. controls. These results suggest that impairments in lexical use are not characteristic of the profile of discourse performances in PD. This is a consistent finding in the existing PD discourse literature (Ash et al., 2011; Ash et al., 2012a; Murray & Lenz, 2001). Researchers reported that the proportion of open class words interacts with syntax

structure in the spoken discourse of individuals with PD. Illes et al. (1988) and Illes (1989) reported that individuals with PD often string together events in a list-like sequence affecting lexical diversity.

No differences were found in the current study in the proportion of verbs spoken between PD vs. control groups. The means and standard deviations suggest that the two groups performed almost equivalently. This is the first study to compare the proportion of verbs in spoken monologic discourse tasks between PD and controls. Holtgraves et al. (2010), using a semi-structured conversation task, went beyond open vs. closed class categories and used Type Token Ratios to calculate the proportion of specific word classes (e.g., verbs, nouns). They reported that individuals with PD with left-side motor symptoms used fewer verbs and fewer closed class words (i.e., function words). Interestingly, the side of motor asymmetry (left vs. right) accounted for 20% of the variance for number of verbs produced and 13% of the variance for number of closed class words produced (Holtgraves & McNamara, 2010). Holtgraves et al. (2010) reported that side of motor asymmetry was the greatest predictor of linguistic performance more so than any of the cognitive tasks administered (i.e., Stroop performance, category fluency, autobiographical memory). These findings suggest that assessing lexical diversity from the perspective of specific word classes (e.g., nouns, verbs) vs. more superordinate categories (e.g., open vs. closed class) may be of benefit in spoken monologic discourse tasks in PD. Additionally, while more research is needed to replicate these findings, Holtgraves et al. (2010) suggests that side of motor symptom asymmetry, in addition to global measures of motor severity and cognition, may be an important variable to consider in discourse performance.

Classifying words using dichotomous categorical labelling systems such as open vs. closed class, while typical in the PD discourse literature, may not be sensitive enough to detect subtle lexical use differences in PD vs. controls. In the current study, the researcher expanded the study of lexical use in PD by including a measure of proportion of verb use. However, the current results suggest that using a broad word-class label such as 'verbs' vs. conducting a microanalysis of verb by specific subtype is not sensitive enough for detecting verb use impairments in PD. For example, analysing the proportion

of dynamic verbs (e.g., process verbs such as *bought, ate*) vs. mental state verbs (e.g., state of being verbs such as *know, like*) in spontaneous language may reveal deficits in verb retrieval relative to the hypothesis that PD-specific disruptions in motor pathways impairs access to semantic knowledge relating to movement features of verbs (Herrera et al., 2012; Pulvermüller, 2005; Rodriguez-Ferreiro et al., 2009). By extension it could be hypothesized that the ratio of dynamic to mental state verbs would differ in the spoken discourse of participants with PD vs. controls. Alternatively, one might examine the occurrence of errors or verbal disruptions relative to the use of specific verb subtypes. Colman et al. (2009) studied verb use in spontaneously generated sentences in a group of Dutch-speaking participants with PD (Colman et al., 2009). Their participants with PD produced significantly more errors on present tense vs. past tense verbs and on intransitive vs. transitive verbs. Similarly, Mayer and Murray (2003) reported that in persons with aphasia, sub-classifying verbs and evaluating the proportion of substantive verbs, excluding auxiliary and modal verbs, was more discriminative than evaluating verbs as a comprehensive word class. Collectively, the existing body of literature underscores the potential complexity of lexical use in PD and the need for researchers to further explore word use with more detailed measures of lexical diversity.

The current study is the first discourse study in PD to collect measures of word retrieval in verbs and to evaluate verb use in discourse. As such it contributes uniquely to the discourse literature in PD. Despite significant group differences in verb retrieval on confrontation naming and fluency tasks, there were no group differences in verb use in spontaneous language. Multiple studies from the aphasia literature have reported paradoxical findings between measures of confrontation naming and word retrieval in spontaneous language (Mayer & Murray, 2003; Pashek & Tompkins, 2002). Possible explanations emerge for this finding in the current study that cannot be fully resolved given the analyses reported herein. Firstly, individuals with PD may have benefited from the added semantic and syntactic contextual support of the discourse tasks in terms of word retrieval of verbs. Secondly, as discussed above, the measurement tool used may have hampered the ability to detect verb retrieval challenges in the spontaneous language of individuals with PD that may manifest as differences in patterns of verb use vs. an overall reduction in proportion of verb use.

Surprisingly, there is a relative paucity of data on lexical diversity in PD using measures that count tokens of word occurrences by specific type (e.g., adjectives, adverbs, transitive verbs, etc.). Recently, research in discourse production in healthy adults and persons with aphasia demonstrated that type-token measures of lexical diversity, even those based on computational and statistical methods (e.g., *D*), are particularly vulnerable to discourse task differences which can affect the reliability of these measures in cross-genre discourse studies (Fergadiotis & Wright, 2011; Fergadiotis et al., 2011). The specific effect of task on lexical diversity in PD has yet to be explored.

Length has been shown to have a powerful effect on lexical diversity (Fergadiotis, Wright, & West, 2013; Heap, 1978). As such, discourse samples of different lengths, as in the Picture and Retell tasks used in the current protocol, can produce very different profiles of lexical diversity if uncorrected methods of lexical quantification are used. Recently, researchers called into question the use of traditional uncorrected lexical diversity measures such as Type-Token Ratios (Templin, 1957) in persons with communication impairments. The work of Fergadiotis et al. (2013) suggests that using statistically adjusted methods of lexical diversity, requiring specialized computer programs, such as the Measure of Textual Lexical Diversity and the Moving-Average Type-Token Ratio produce more accurate and unbiased estimates of lexical diversity in the spoken language of individuals with aphasia. In the current body of PD discourse literature, including the current study, the conclusions drawn relative to lexical diversity impairments in PD discourse may be biased by the measures used to assess lexical use in PD. These measures, largely quantitative in nature, may overlook qualitative changes in lexical use. While this area of research warrants further exploration in PD, the recent work of Fergadiotis et al. (2013) is a reminder that researchers should be rigorous in their selection of measures for lexical diversity.

Verbal Fragmentation. The occurrence of verbal disruptions (i.e., word repetitions, phrase revisions, pauses, word choice or sound errors) can signal a break down in on-line processes of cognition, language, and/or motor planning/execution

during spontaneous spoken language. Verbal disruptions occur even in the spoken language of healthy adults and in normal proportions do not disrupt the productivity or flow of information (Roberts et al., 2009). However, in the presence of acquired or degenerative disorders the prevalence and location of verbal disruptions can significantly affect spoken language (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001; Davis & Maclagan, 2010; Jokel et al., 2007). The study of verbal fragmentation is of great interest in PD. Of particular interest is elucidating whether episodes of verbal disruption are related to motor speech errors or to cognitive-linguistic errors. In the current study, an expanded definition of verbal disruptions was applied in order to capture a larger variety of both word (i.e., revisions, repetitions, incomplete utterances, paraphasic errors) and non-word (i.e., long pauses, interjections such as ‘uh’ and ‘uhm’) classes of verbal disruptions.

Verbal fragmentation accounted for almost 20% of the variance associated with group differences and demonstrated a larger effect size than productivity relative to discriminating PD vs. control groups. Individuals with PD were more impaired vs. controls on the number of disruptions per 100 words whereas total disruptions and disruptions per minute did not significantly differ from controls. In other words, individuals with PD did not differ from controls in the total number of disruptions but in the density of verbal disruptions. These findings are consistent with those of multiple researchers who have observed both quantitative and qualitative differences in patterns of verbal disruption in PD and controls (Illes et al., 1988; Illes, 1989; McNamara et al., 1992). However, not all findings are congruent among studies. Huber and Darling (2011) reported no differences between PD vs. controls in proportion of total disfluencies during a spoken discourse task. Ash et al. (2012a) also found no PD vs. control differences on a composite variable of more ‘linguistic’ verbal disruptions (i.e., fillers, repetitions, reformulations) but did find significant differences in the proportion of unfilled pauses (i.e., > 2 seconds) between groups.

Ash et al. (2012a) reported a composite variable for verbal disruptions that included phonemic errors, phonetic errors, repetitions, revisions, and abandoned utterances. They referred to their global measure of verbal disruption, as ‘articulation errors’ (Ash et al.,

2012a). A key methodological difference between the current study and the Ash et al. (2012a) study is that in the current study the total number of verbal disruptions included unfilled pauses exceeding two seconds duration which were analysed as a separate category in the Ash et al. study. Ash et al. (2012a) reported no differences between PD and controls for articulation errors/100 words. However, differences emerged for the parkinsonism group with dementia. As with other findings from the Ash et al. studies (2011; 2012a), using their data to interpret the effect of cognition on discourse in PD is challenging given that their group with cognitive decline was a mixed dementia group including participants with PDD and DLB. While these two disorders share partially overlapping pathologies and are considered part of a single spectrum of disorders, they are different dementia subtypes (Lippa et al., 2007; McKeith & Mosimann, 2004). Moreover, while individuals with DLB have parkinsonism, they do not have PD (Lippa et al., 2007). For example, individuals with DLB may not have the same motor speech profile or severity as individuals with PD (Müller et al., 2001). Furthermore, research exploring potential cognitive-linguistic impairments in DLB is sparse compared to PD. As such, the mixing of these two dementia subtypes presents challenges in the interpretation of the additive effect of cognitive impairment in PD on spoken discourse performance, particularly for verbal disruptions, which are highly sensitive to disruptions in both motor and cognitive-linguistic processes.

Verbal disruptions interact with task demands, in addition to other discourse variables. In the current study, significant moderate-sized task effects (without interactions) were observed for the total number of disruptions, suggesting that the unique demands associated with the Retell task increased the prevalence of verbal disruptions. Verbal disruptions can interact with multiple domains of discourse performance. For example, productivity measures can be adversely affected by verbal disruptions leading to an increase in the duration of discourse samples without increasing the total words or efficiency of output. Additionally, verbal disruptions such as semantic paraphasias and empty words, resulting from potential challenges in word retrieval, can interact with the proportion of CIUs and CIU rate measures.

As expected, a large diversity of verbal disruption subtypes occurred in the discourse of both participants with PD and controls. However, verbal fragmentation for the PD group differed from controls not only in the density of verbal disruptions (i.e., Disrupt/100 Words) but also in the quality of verbal disruptions. While both participants with PD and controls exhibited phrase repetitions and phrase revisions/reformulations, typically the word-level verbal disruptions exhibited by individuals with PD were more complex than those exhibited by controls. The presence of semantic and phonemic paraphasias, while occurring infrequently in the current study, was unique to the PD group.

Word and phrase level repetitions can occur commonly in the discourse of healthy adults. They can signal challenges in the planning and or execution of word choices, grammar choices, or even changes in thought direction (Levelt, 1999; Merlo & Mansur, 2004; Roberts et al., 2009). In healthy adult speakers higher frequency disruptions include interjections (i.e., *uh, uhm*) and revisions/reformulations (i.e., */he said/ no she said it is time*) (Roberts et al., 2009). When occurring in lower proportions (< 10 per 100 intended syllables) (Roberts et al., 2009), they are considered a component of normal spontaneous language performance. In contrast, repetitions of more than one word (i.e., *he went he went to the store*), repetitions of multisyllabic words (i.e., *uni-univer-university-university*), and sound/syllable repetitions (i.e., *she f-f-fell*) are less typical in healthy adult speakers (Roberts et al., 2009) and mark challenges in motor and/or language planning/execution in on-line language production. Less typical disfluencies, including semantic and verbal paraphasias, while not occurring frequently occurred exclusively in the PD group. Also, complex verbal disruptions comprised of multiple subtypes of disruptions (i.e., *I-I-I sssaw the girl, no the boy, the girl pointing for some di-directions, showing some directions*) were observed predominately in the PD group. These observations suggest that verbal disruptions between PD vs. controls differed not only quantitatively but also qualitatively.

Summary. Cross-study comparisons of verbal disruptions are complicated by the fact that among discourse studies in PD, a variety of different verbal disruptions constructs are applied. With the exception of Ash et al. (2012a), the majority of researchers in PD discourse have taken a less comprehensive approach focussing on a

specific aspect of verbal fragmentation (e.g., pausing) or a specific set of verbal disruption subtypes (e.g., within word vs. between word). Notwithstanding the different approaches, multiple researchers reported that the frequency and quality of verbal fragmentation distinguishes the discourse of individuals with PD from that of controls (Ash et al., 2012a; Illes et al., 1988; Illes, 1989; McNamara et al., 1992). Despite the large effect size associated with the canonical variable, only a moderate effect size was found for the variable Disrupt/100 words. While the intent was to obtain a comprehensive profile of verbal disruptions, it is possible that the inclusive reporting of verbal disruptions in the current study occluded the ability to detect a single verbal disruption subtype (e.g., pauses or sound repetitions) that may be more discriminative than the global variable.

Informativeness. In keeping with the existing literature in PD reduced informativeness, in the form of reduced overall amount and efficiency of information features prominently in the characterization of discourse impairments in PD (Bayles, 1990; Godbout & Doyon, 2000; McNamara et al., 1992; Murray, 2000). In the current study, informativeness was the most distinguishing variable among all those assessed accounting for 45.1% of the variance associated with group differences. Furthermore, each of the three dependent variables comprising the canonical variable Informativeness (i.e., Total CIUs, % CIUs, and CIUs per minute) differed significantly from controls with observed effects of moderate to large magnitudes. While measures of CIUs are common in the discourse literature in healthy adults and other disorders, only one other previous study, Murray (2000), included CIU measures (specifically % CIUs) in its analyses of PD discourse. The results of the current study reinforce those of Murray (2000) who reported significant differences between PD and controls for % CIUs using a picture description task. In the current study, we extended Murray's findings by expanding the analysis to multiple discourse tasks and including CIU measures that reflect volume, density, and efficiency of informativeness. The discriminant function analysis showed that each of these variables contributed unique information to the discrimination of discourse samples produced by PD vs. control groups. Importantly and perhaps surprising given that PD is a motor disease, the discriminant function weightings indicate that impairments in informativeness, as measured by CIUs, feature more prominently

than impairments in productivity in the characterization of discourse performance in PD. The construct of words vs. CIUs may offer a potential explanation for why CIU measures showed larger effects vs. productivity measures. All linguistic units meeting the definition of a word are included in productivity measures. However, to meet the definition of a CIU words must be correct to the context of the picture and represent novel information. Repeated words not used as a linguistic device for emphasis or not adding new information do not count in CIU informativeness measures but would count in productivity measures. As such, CIU measures are sensitive to word retrieval errors and word/phrase level verbal disruptions. Therefore, CIU measures, more than productivity measures are affected by the integration of cognitive, language, and motor aspects of spoken language. Given the nature of the impairments in PD, it is not surprising that CIU measures feature so prominently in the profile of discourse impairment. While studies of informativeness in PD spoken discourse are rare, the current study suggests that it is an important, if not the most important, variable in the profile of discourse impairment in PD. As such, it warrants a much greater consideration in the PD discourse literature.

Summary. Conducting a multi-level discourse analyses protocol using a multi-genre sampling approach, the researcher was able to discern a comprehensive profile of discourse impairment in PD that included impairments in productivity, grammar, informativeness, and verbal fragmentation. Notwithstanding the challenges in comparing findings across studies due to differences in participant inclusion/exclusion criteria and methodology, the profile that emerged from the current study aligns generally with the existing body of discourse literature in PD. However, there are some important areas of divergence. The most notable points of divergence are in the domains of productivity (specifically total words) and in verbal fragmentation (specifically the proportion of verbal disruptions per 100 words). The PD participants and controls in the current study were fairly consistent with those reported in other studies. Motor severity was slightly higher for the PD participants in the current study. However, the issues associated with reliability of the UPDRS-III may inflate these perceived differences. Otherwise there were no egregious differences in the participant groups among studies that would explain the observed differences. As such, it is more likely that these discrepancies result from

differences in language sampling techniques. One of the goals of the current study was to collect a robust and representative discourse sample of sufficient size and representativeness for optimizing the discourse analyses toward the objective of creating a profile of discourse impairment in PD. The published work on discourse methodology would suggest that the differences observed between the current study and the existing literature largely result from the sampling approach that revealed patterns of performances not readily discernable using techniques reported in previous studies.

RQ 4: To what degree does a unique profile characterizing discourse impairments in PD discriminate the spoken language of participants with PD from that of controls?

The analyses conducted via the multiple mixed MANOVAs enabled the researcher to elucidate the measures of discourse performance that differed significantly between groups. However, the resulting list of affected variables was only a partial characterization of discourse impairment in PD in that it assumed an equal contribution of each of the identified domains of discourse impairment to the characterization of discourse in PD. To refine and to optimize the profile of discourse impairment in PD, a second level of analysis was conducted to determine the ideal weighting of each variable, in the context of the remaining variables. The discriminant function analysis was used to create an optimized weighting of discourse variables that could separate effectively discourse samples produced by individuals with PD from those produced by controls. Once the optimized profile of discourse impairment in PD was created, it was tested to determine its effectiveness for categorizing discourse samples produced from two different discourse genres (Picture vs. Retell) into either PD vs. control groups. The assumption was that the more accurate and the more complete the profile of discourse impairment, the better the discriminant function would be at separating the discourse samples. Relative to **RQ 4**, the resulting product was a mathematical function that when applied to a limited set of discourse data generated a composite score for discourse performance that was used not only to classify existing discourse samples with > 80% sensitivity and specificity but could be used to classify new discourse samples not

included in the existing study data. This is a unique approach to profiling discourse impairments that has not been previously reported in the PD discourse literature.

Based on the MANOVA analyses, the discourse profile that was assessed in the discriminate function analysis included the following variables: total words, words per minute, percentage of grammatical C-units, CIUs, % CIUs, CIUs per minute, and number of verbal disruptions per 100 words. While the productivity measure variable Total Words was identified as a discourse measure that significantly differed between PD vs. controls, it was ultimately removed from the discourse function. The contribution of the variable Total Words to the model was redundant with the variable CIUs. In other words, while PD vs. controls may have differed in total words produced, this marker of productivity impairment in PD was not a sufficient source of unique information to the characterization profile. Intuitively, one would think that CIUs, % CIUs and CIUs per minute would provide redundant information. However, the discriminate function analysis showed that each of these variables contributed unique information to the discourse function without significant overlap. Overall, the weighting of variables in the discourse function reinforces the findings of the MANOVAs in that informativeness and verbal fragmentation variables are more heavily weighted in the discriminant function equation.

The final discourse algorithm included WPM, percentage of grammatical C-units, CIUs, % CIUs, CIUs per minute and number of verbal disruptions per 100 words. The discourse discriminant function successfully classified a total of 76 discourse samples (38 Picture and 38 Retell) with a sensitivity of 81.6% and a specificity of 86.8%. These results suggest that the discourse function was correctly classified 81.6% of PD discourse samples into the PD group. Furthermore, it indicates that the function correctly classified 86.8% discourse samples that were not generated by a participant with PD as not belonging to that group. A visual analysis of the data indicates that overall the discourse samples elicited via the Picture task were classified correctly more often than those elicited with the Retell task. Although this work is still in preliminary stages, it suggests that while cross-genre sampling was essential for developing the profile of discourse impairment, clinically the Picture task may be more useful given its greater classification

accuracy. One of the more interesting nuanced findings is that while the initial profile of discourse impairment was derived from data that were averaged across the two discourse tasks, the resulting algorithm was able to correctly classify individual discourse samples elicited using either the Picture or Retell task individually.

Summary. Certainly, these findings do not imply that PD can be diagnosed using spoken discourse. However, the effectiveness of the discourse function in categorizing the spontaneous language samples produced by PD vs. controls, at a preliminary level, strengthens the validity of and verifies the accuracy of the profile of discourse impairment that the researcher presents here. Importantly, this work creates a foundation for further testing and refining of the algorithm in the context of PD, PD-MCI and PDD. While discourse performance is not essential to the diagnosis of PD, it may be valuable in discerning subtle, early changes in cognition and language that can be harbingers of dementia in PD. Although the discourse function presented here is preliminary and requires further testing, it may have significant future research and clinical relevance.

RQ 5: What is the nature of and strength of the relationship between discourse performance and markers of disease severity in PD, age, education, and global cognitive function?

Previous studies of cognition and language in PD elucidated the association among performances in these domains and markers of disease severity. Relative to **RQ 5**, the significant relationships were such that more impaired discourse performances correlated with worsening disease severity and worsening global cognition. Discourse performances on the Picture task shared a stronger relationship with indicators of disease severity vs. the Retell task. Age and education did not correlate with discourse performance, with the exception of education and Retell performances.

In the current study, discourse performance as measured by the discriminant function score on the Picture task was significantly and negatively associated with LED, Hoehn and Yahr scores, and UPDRS-III scores. The associations indicate that as disease severity worsened, discourse performance also worsened. The discriminant function score was not associated with duration of disease. However, for the Retell task there

were no significant associations between the scores from the discourse discriminant function and any of the measures of disease severity. This finding reinforces the point made earlier that the Picture task may be of greater clinical utility not only for its better classification abilities but also because of its significant association with markers of disease severity.

Global cognitive function was associated strongly and positively with discourse performance for both tasks. This finding suggests that better cognition, even in the absence of dementia and MCI, was associated with better discourse performance. This was an expected finding that is consistent with reports from the discourse literature in PD and healthy aging (Cannizzaro & Coelho, 2013; Murray, 2000). While age and education were not significantly correlated with discourse performance in the Picture task, education was associated with the Retell task. Higher education was associated with better discourse performance on the Retell task. It is possible that this reflects the influence of the verbal memory demands of the Retell task on discourse performance such that individuals with higher levels of education presented with better verbal memory skills and therefore produced discourse samples that were not as affected by the intrinsic demands of the task. This is consistent with the work of Ardila, Ostrosky-Solis, Rosselli, and Gómez (2000) who reported that education has a protective effect from the typical age associated declines observed in verbal memory. Given that discourse studies in healthy aging and disease have reported significant relationships among verbal episodic memory and a number of discourse variables (Ash et al., 2012a; Light & Anderson, 1985; Murray, 2000; Peach, 2013) this is a plausible explanation for the task differences observed for the relationship between education and discourse performance in the current study. The lack of association between age and discourse performance is not of great surprise in the current study in that multiple researchers reported the relative stability microlinguistic and macrolinguistic measures of discourse in the context of spontaneous language changes in older adults (Capilouto et al., 2005; Marini et al., 2005; North, Ulatowska, Macaluso-Haynes, & Bell, 1986; Obler et al., 1994).

While previous discourse studies in PD have not used a single discourse score, the finding that discourse performance (or aspects of discourse performance) are associated

with measures of disease severity and cognition are generally consistent with the existing body of PD discourse literature. Ash et al. (2012a) reported significant, positive correlations between WPM and measures of cognition. However, they reported no association between WPM and motor severity (subset of items from the UPDRS-III) (Ash et al., 2012a). Illes et al. (1988) reported strong correlations between syntactic complexity and disease severity such that more severe disease was associated with less complex syntax in spoken language. In her study, Murray (2000) reported significant and very strong (.866) positive associations between % CIUs and global cognition. Given the weighting of informativeness variables in the discriminant function, Murray's findings suggest that the association between discourse performance and global cognition observed in the current study may be driven largely by the influence of the informativeness variables (CIUs, % CIUs, CIUs per minute). In the Murray study, no other significant correlations were reported between global cognition and other macrostructural measures of discourse. However, Murray and Lenz (2001), using a personal narrative task, reported significant relationships between a global measure of cognition and multiple measures of syntax complexity.

Relative to disease severity, Murray (2000) assessed the relationship between years diagnosed with PD and discourse performance across a variety of individual variables. The only significant relationship was reported for total utterances such that as years with disease increased total utterances, a marker of productivity decreased. In keeping with the current study, the relationship between duration of disease and discourse performance is not as strong as for other markers of disease severity. While typically longer disease duration is associated with more severe motor symptoms, the existence of potentially different subtypes of phenotypic presentation in PD complicates this picture somewhat (Thenganatt & Jankovic, 2014). There is a growing body of evidence to support the subtyping and separate characterization of PD phenotypes (Thenganatt & Jankovic, 2014). For example, individuals with tremor-dominant PD vs. those with rigidity and gait dysfunction progress more slowly and experience a lesser degree of cognitive decline. While the sample size in the current study was not sufficiently powered for a subgroup analysis, further work exploring the relationship between discourse impairment and disease phenotype is indicated. Other researchers reported that speech intelligibility

correlates significantly with syntax and that improved speech intelligibility is associated with more complex and more accurate syntax production in spoken language (Illes et al., 1988; Murray & Lenz, 2001). Whether the results of the correlation analyses reflect a relationship specifically between motor speech performance and grammar or reflect general motor decline and disease progression in PD remains unclear.

Summary. These findings suggest that there is a relationship between discourse performance and global cognition in addition to a relationship between discourse performance and disease severity. However the exact nature of these relationships relative to the pathology of PD remains ambiguous. Collectively, the current study in the context of the existing literature in PD discourse underscores the importance of increasing our understanding of the relationships among motor severity, cognition and advancing disease processes in PD on spoken language performance from both a clinical and research perspective.

5.4 Study Limitations

A first limitation is that the protocol did not include a comprehensive or psychometrically derived screening measure for depression. Given the high prevalence of depression in PD and the relationship between depression and cognitive performance in both PD and healthy adults, the current protocol would have benefited from a more rigorous depression screening process. Only Murray (2000) and Murray and Lenz (2001) have previously used formal depression screening measures in the study of PD and discourse performance. However, they did not correlate the depression measure with any of the discourse measures. Therefore, while depression and anxiety have been shown to affect cognitive performances in PD, there is currently no existing evidence from which to determine if spoken discourse performance is impaired in PD as a function of depression or anxiety. However, recently Murray (2010) using tasks and discourse measures similar to those reported here, explored whether or not spoken discourse performance differed among healthy older adults with and without depression and individuals with Alzheimer's disease. She reported no significant differences between older adults with and without depression for spoken discourse measures assessing productivity,

informativeness, and grammar. Moreover, there were no correlations among scores on the Geriatric Depression Screen and measures of discourse performance. While the author used self-report, a cursory screening measure and a medical chart review for individuals with PD, these measures likely were not sufficiently robust to confidently exclude depression or anxiety effects on discourse performance given the particular screening measure used (Schrag et al., 2007) and issues with under identification and under reporting of depression in clinical settings (Shulman, Taback, Rabinstein, & Weiner, 2002). Furthermore, using a formal tool in the control group may have been more reliable for identification of depression than a clinical history interview. However, one of the complexities of using formal screening measures for depression in PD is that the symptoms of the disease itself can often overlap with depressive symptoms, which may inflate scores on depression screening measures (Marsh, McDonald, Cummings, & Ravina, 2006). As such, careful attention to the selection and interpretation of depression screening tools must be considered in the context of research in PD discourse. Notwithstanding the importance of depression and anxiety on cognitive performance, the existing literature in both older adults and PD, does not suggest that in the current study the presence of depression/anxiety would have significantly accounted for the findings vs. the effects of PD.

Another potential limitation is that the protocol did not include a direct measure of IQ or an estimated IQ score. Murray and Lenz (2001) reported an estimated IQ score in their PD discourse study (Murray & Lenz, 2001). While measures of intelligence, particularly verbal intelligence, are linked to discourse performance in children, the relationship in adults is less clear (Cannizzaro & Coelho, 2013). Notwithstanding the ambiguous findings in the literature, ensuring that participants were comparable on a measure of verbal IQ, especially given the large range of years of education in the current study may have benefited the interpretation of the results.

Thirdly, the value of the discriminant function developed in the current study is limited to the analysis of discourse samples using the picture sequence description and story retell tasks. Furthermore, presently the profile of discourse impairment that has been identified is limited to individuals with PD with similar motor and cognitive profiles to the current

participants. While the SPSS protocol for this procedure does conduct a cross-validation using a leave-one-out analysis, the discriminant function has not been cross-validated using an orthogonal sample. The current study assessed individuals with PD in an optimal 'on' medication state, as such the discourse profile reported here is that of PD in the context of pharmacological treatment for the symptoms of PD and not strictly PD in isolation. This work, while promising, is preliminary and cannot be confidently generalized to the classification of discourse samples elicited via other stimuli or elicitation tasks. Moreover, it requires further testing before it can be applied to other cohorts of individuals with PD and those with concomitant dementia. It is possible and indeed likely that other discourse variables not included in the current discriminant function for profiling discourse in PD (e.g., main event analyses, global coherence) may play a role in a discriminant function designed to distinguish PDD.

Another potential limitation is the statistical power for the current study protocol. The current study was powered to detect study effects of a medium to large magnitude. As such, more subtle discourse impairments may have been obscured. In the current study a more liberal significance testing criteria was applied to optimize the power of the study given the statistical tests used and the available resources. While the power of the study may have been slightly compromised particularly for interaction effects, with the exception of Ash et al. (2011) and Ash et al. (2012a), the *N* size reported herein well exceeds that published in the existing PD discourse literature. Although the power assumed for the study may be considered a potential limitation in the interpretation and generalizability of these findings, it is important to note that with the exception of total words and disruptions/100 words all of the variables comprising the profile of discourse impairment in PD were significant for group effects at a .05 level of significance and many were significant at an alpha level of .001. This suggests that while increasing the sample size of each group may have resulted in more pronounced findings for select measures of productivity and verbal fragmentation, a larger *N* size may not have fundamentally changed the profile of discourse impairment that emerged from the results.

5.5 Theoretical and Clinical Relevance of the Findings

While the work presented herein is preliminary, it has important theoretical and clinical implications. The current work advances models of discourse performance in adults with degenerative neurological disorders by demonstrating that cognitive, language, and motor disruptions are uniquely manifested in spoken monologic discourse tasks across multiple domains of discourse performance. These findings reinforce the importance of spoken discourse tasks in elucidating subtle, yet important, changes in cognition, language, and motor aspects of communication. Importantly, using robust sampling practices, the current study resolves ambiguities in the current PD discourse literature confirming that spoken discourse in PD is characterized by impairments in productivity, grammar, informativeness, and verbal disruption. With a better understanding of the model of spoken discourse in PD, researchers can move forward to further our understanding of spoken language impairments in PD with and without dementia and inform our understanding of the potential roles of dopaminergic systems and subcortical structures in language production.

Furthermore, these findings have important theoretical implications in PD. Researchers have typically explored discrete aspects of language performances from a perspective of better understanding cognitive impairments in PD. Similarly, researchers have explored connected spoken language tasks but typically through the lens of motor speech disruptions in PD. However, this study in the context of the existing discourse literature of spoken discourse impairments in PD, confirms the presence of spoken *language* impairments in PD that are unique from the motor speech consequences of PD. As such, the current study helps inform the findings reported by Miller and colleagues relative to the nature of communication challenges in PD (Miller et al., 2006) by validating the perceptions of individuals with PD that their communication challenges are not limited to changes in speech and voice. This has important clinical implications in terms of targets and models of treatment. The current findings suggest with certainty that a reconsideration of our current models of speech-language pathology interventions in PD is required and should reflect the dynamic nature of communication challenges in PD. Specifically, the finding that the spoken language of individuals with PD differs from

controls in a predictable and characteristic fashion suggests that social and participation approaches which allow clinicians to address cognition, language, and motor targets within an integrated therapy paradigm that focuses on communication and not just motor speech or voice as isolated targets are important to consider in PD rehabilitation.

The methodological approach in the current paper furthers the existing literature in PD and also the spoken discourse literature in general. These findings confirm the benefit of robust sampling approaches both in terms of size of samples and representativeness of discourse samples. The current findings suggest that many of the discrepancies observed among the existing body of PD discourse studies can be attributed to sampling practices (i.e., effects of sample size and task differences) vs. participant variability. The current sampling practices and statistical approaches enabled the development of a single unified function of discourse in PD. This approach is distinctive in the existing PD literature and unique in the discourse literature in general. As such, the current study presents a novel methodology for understanding discourse performances in PD and in other disorders.

Finally, the current study advances our understanding of assessment in PD and serves as a foundation for improving assessment practices in PD with concomitant dementia. While protocols for examining cognitive (i.e., particularly frontal lobe functions) are well established in the PD literature, formal assessment of language is typically limited to a measure of confrontation naming. In the field of communication sciences and disorders, there is only one test of cognitive-linguistic function with PD specific norms. The Arizona Battery of Communication Disorders of Dementia (ABCD) has both PD and PDD norms based on a small normative sample (Bayles & Tomoeda, 1993). Our knowledge of cognitive-linguistic deficits in PD has greatly expanded since publication of the ABCD highlighting the need to develop other PD-specific measures.

The current protocol proposes a preliminary model for the assessment of spoken language in PD that can be further refined toward this end. The finding of significant differences on two measures of word retrieval for verbs and marginal statistical significance for a test of semantic knowledge reinforces the critical need to advance our current standardized and non-standardized assessment practices in PD developing rigorously designed

psychometric measures that reflect the current evidence regarding language and communication challenges in PD complementing them with measures of spoken discourse.

5.6 Future Directions

The next stage of this work is to refine the profile of discourse impairment in PD to determine whether the sensitivity and specificity of the discriminant function can be optimized not only for individuals with PD but also for populations of PD patients with concomitant MCI and dementia processes. Conducting another layer of analyses that assesses specific subtypes of verbal fragmentation behaviours, measures of global coherence and maintenance, and main event unit analyses will be useful for revealing other aspects of discourse abilities that would distinguish the performances among these groups.

A better understanding of the profile of specific verbal disruptions in PD will help identify a specific disruption subtype that is potentially more effective in discriminating the discourse of PD from controls than the global measure of verbal disruptions used in the current study. Furthermore, improving our understanding of the relationship between verbal disruptions and other aspects of discourse (e.g., grammar) will have important clinical implications for PD.

Several researchers reported the value of main event unit analyses in understanding patterns of spoken discourse performance of individuals with aphasia and healthy older adults (Brookshire & Nicholas, 1984; Capilouto et al., 2005). Previously in the PD literature, Murray (2000) reported no significant differences in the proportion of informative utterances between PD and controls. Moreover, Bayles (1990) reported no differences in the number of main ideas expressed between individuals with PD who did not have dementia and healthy older adults using a story-retelling paradigm.

Notwithstanding the importance of the existing findings, establishing the relevance of main event analyses to the discriminative ability of a profile of PD discourse using rigorous discourse sampling methods is warranted. Once the protocol is fully optimized it can be tested across a variety of discourse tasks and cross-validated orthogonally to

assess its discriminative effectiveness among a larger range of discourse elicitation paradigms and the populations of individuals with PD.

One of the greatest potentials for this work is in assessing how the discriminant function can be used to inform profiles of and diagnostic accuracy of cognitive impairment in PD either at the level of MCI or dementia. Using a foundation built upon a strong discriminant function for PD without dementia, a profile of discourse impairment can be established for both PD-MCI and PDD using approaches similar to those described in the current study. The value of discourse in the detection of early cognitive and language changes in other dementias has already been reported (Fleming & Harris, 2008). Importantly, spoken discourse tasks can be sensitive measures for detecting longitudinal changes in cognition and language in dementia and other degenerative neurological disorders (Kemper, Thompson, & Marquis, 2001; Roberts-South et al., 2012; Tomoeda & Bayles, 1993). Aside from improving the sensitivity for both diagnostic and discriminative purposes, a better understanding of the profiles of discourse challenges in PD may help inform and advance therapeutic options in PD, PDD, and PD-MCI. As such, using the discourse discriminant function to establish hierarchies of task difficulties as well as assessing the test-retest reliability of the function for its use as a longitudinal change measure is an important consideration for future work.

Furthermore, future studies should explore in greater detail the relationship between spoken discourse impairments and variables specific to PD. This area of research would benefit tremendously from understanding changes in spoken language relative to ‘on’ vs. ‘off’ medication states, side of motor asymmetry, and emerging subtypes of PD. Such studies may further inform not only classification approaches in PD but also a deeper understanding of the roles of dopaminergic and subcortical systems in cognition and language.

Finally, and perhaps most importantly, there is a need to determine to what degree a profile of discourse impairment corresponds to everyday challenges of communication and quality of life that are reported by individuals with PD and their care partners. Previous studies showed that performance on more structured discourse tasks correlate

with conversation performance observed in everyday activities (Doyle et al., 1995). However, this work does not presently exist in the PD literature. Future studies should investigate the ecological validity of spoken monologic discourse tasks in relation to conversation and every day communication challenges in PD. In isolation, establishing a profile of discourse impairment in PD is both interesting and informative from a clinical and research perspective. However, the development of a ‘clinic friendly’ discourse profiling measure that could be used to both assess and guide clinical practice in PD, PD-MCI and PDD would be the optimal goal. To this end, there is a need for research to evaluate the relationship between PD discourse performance and measures of communication quality of life, health related quality of life, and conversation performance in everyday contexts.

5.7 Conclusion

The current study adds clarity to the existing body of PD discourse literature and advances significantly our understanding of spontaneous spoken language impairments in PD. PD participants’ patterns of cognitive and language performances on a comprehensive battery of standardized measures align with a model of spoken discourse abilities that emerges from the data. There is preliminary evidence to establish a characteristic profile of spontaneous language impairments in PD that is distinguishable from controls and is uniquely revealed through performances on spoken monologic discourse tasks using a cross-genre, multi-level discourse analysis paradigm. This profile, expressed as a single discriminant function for discourse performance, is strongly associated with global cognitive abilities. In PD, the discriminant function applied to the picture sequence task in relation to Hoehn and Yahr scores, UPDRS-III scores, and levodopa equivalent dose suggests that there is a relationship between discourse performance and motor severity and disease staging such that performances in spoken language declines with increased disease/motor severity. This preliminary work has important theoretical and clinical implications relative to our understanding of the roles of dopaminergic and subcortical systems in language production and also relative to models of speech-language pathology services in PD.

References

- Aarsland, D., Andersen, K., Larsen, J. P., & Lolk, A. (2003). Prevalence and characteristics of dementia in Parkinson disease: an 8-year prospective study. *Archives of Neurology*, 60(3), 387-392.
- Aarsland, D., Beyer, M. K., & Kurz, M. W. (2008). Dementia in Parkinson's disease. *Current Opinion in Neurology*, 21(6), 676-682. doi:10.1097/WCO.0b013e3283168df0
- Aarsland, D., Brønnick, K., & Fladby, T. (2011). Mild cognitive impairment in Parkinson's disease. *Current Neurology and Neuroscience Reports*, 11(4), 371-378.
- Aarsland, D., Brønnick, K., Williams-Gray, C., Weintraub, D., Marder, K., Kulisevsky, J., . . . Emre, M. (2010). Mild cognitive impairment in Parkinson disease: a multicenter pooled analysis. *Neurology*, 75(12), 1062-1069. doi:10.1212/WNL.0b013e3181f39d0e
- Aarsland, D., Perry, R., Brown, A., Larsen, J. P., & Ballard, C. (2005). Neuropathology of dementia in Parkinson's disease: a prospective, community-based study. *Annals of Neurology*, 58(5), 773-776. doi:10.1002/ana.20635
- Alexander, G. E., & Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends in Neuroscience*, 13(7), 266-271.
- Alexander, G. E., Crutcher, M. D., & DeLong, M. R. (1990). Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. *Progressive Brain Research*, 85, 119-146.
- Alpert, M., Rosen, A., Welkowitz, J., & Lieberman, A. (1990). Interpersonal communication in the context of dementia. *Journal of Communication Disorders*, 23(4-5), 337-346.
- American Speech-Language-Hearing Association. (1997). Guidelines for audiologic screening.
- Araujo, N. B., Barca, M. L., Engedal, K., Coutinho, E. S., Deslandes, A. C., & Laks, J. (2011). Verbal fluency in Alzheimer's disease, Parkinson's disease, and major depression. *Clinics (Sao Paulo)*, 66(4), 623-627.
- Arbuthnott, K., & Frank, J. (2000). Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *Journal of Clinical and Experimental Neuropsychology*, 22(4), 518-528. doi:10.1076/1380-3395(200008)22:4;1-0;FT518
- Ardila, A., Ostrosky-Solis, F., Rosselli, M., & Gómez, C. (2000). Age-related cognitive decline during normal aging: the complex effect of education. *Archives of Clinical Neuropsychology*, 15(6), 495-513.
- Ash, S., McMillan, C., Gross, R. G., Cook, P., Gunawardena, D., Morgan, B., . . . Grossman, M. (2012a). Impairments of speech fluency in Lewy body spectrum disorder. *Brain and Language*, 120(3), 290-302. doi:10.1016/j.bandl.2011.09.004
- Ash, S., McMillan, C., Gross, R. G., Cook, P., Morgan, B., Boller, A., . . . Grossman, M. (2011). The organization of narrative discourse in Lewy body spectrum disorder.

- Brain and Language*, 119(1), 30-41. doi:10.1016/j.bandl.2011.05.006
- Ash, S., McMillan, C., Gunawardena, D., Avants, B., Morgan, B., Khan, A., . . . Grossman, M. (2010). Speech errors in progressive non-fluent aphasia. *Brain and Language*, 113(1), 13-20. doi:10.1016/j.bandl.2009.12.001
- Ash, S., Moore, P., Antani, S., McCawley, G., Work, M., & Grossman, M. (2006). Trying to tell a tale: discourse impairments in progressive aphasia and frontotemporal dementia. *Neurology*, 66(9), 1405-1413. doi:10.1212/01.wnl.0000210435.72614.38
- Ash, S., Xie, S. X., Gross, R. G., Dreyfuss, M., Boller, A., Camp, E., . . . Grossman, M. (2012b). The organization and anatomy of narrative comprehension and expression in Lewy body spectrum disorders. *Neuropsychology*, 26(3), 368-384. doi:10.1037/a0027115
- Barbas, H., Garcia-Cabezas, M. A., & Zikopoulos, B. (2013). Frontal-thalamic circuits associated with language. *Brain and Language*, 126(1), 49-61. doi:10.1016/j.bandl.2012.10.001
- Bartels, A. L., & Leenders, K. L. (2009). Parkinson's disease: the syndrome, the pathogenesis and pathophysiology. *Cortex*, 45(8), 915-921. doi:10.1016/j.cortex.2008.11.010
- Bayles, K. (1990). Language and Parkinson disease. *Alzheimer Disease and Associated Disorders*, 4(3), 171-180.
- Bayles, K., & Tomoeda, C. (1993). *The Arizona Battery for Communication Disorders of Dementia*. Austin, TX: Pro-ed.
- Benito-Leon, J., Louis, E. D., Posada, I. J., Sanchez-Ferro, A., Trincado, R., Villarejo, A., . . . Bermejo-Pareja, F. (2011). Population-based case-control study of cognitive function in early Parkinson's disease (NEDICES). *Journal of Neurological Sciences*, 310(1-2), 176-182. doi:10.1016/j.jns.2011.06.054
- Berg, E., Bjornram, C., Hartelius, L., Laakso, K., & Johnels, B. (2003). High-level language difficulties in Parkinson's disease. *Clinical Linguistics and Phonetics*, 17(1), 63-80.
- Bertella, L., Albani, G., Greco, E., Priano, L., Mauro, A., Marchi, S., . . . Semenza, C. (2002). Noun verb dissociation in Parkinson's disease. *Brain and Cognition*, 48(2-3), 277-280.
- Bhat, S., Iyengar, K. R., & Chengappa, S. (2001). Pragmatic deficits in Parkinson's disease: Description of two case studies. *Journal of the Indian Speech & Hearing Association*, 15, 79-84.
- Bohlhalter, S., Abela, E., Weniger, D., & Weder, B. (2009). Impaired verbal memory in Parkinson disease: relationship to prefrontal dysfunction and somatosensory discrimination. *Behavioral and Brain Functions*, 5, 49. doi:10.1186/1744-9081-5-49
- Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech*, 44(2), 123-147.
- Boulenger, V., Mechtouff, L., Thobois, S., Broussolle, E., Jeannerod, M., & Nazir, T. A. (2008). Word processing in Parkinson's disease is impaired for action verbs but not for concrete nouns. *Neuropsychologia*, 46(2), 743-756. doi:10.1016/j.neuropsychologia.2007.10.007

- Braak, H., & Del Tredici, K. (2008). Cortico-basal ganglia-cortical circuitry in Parkinson's disease reconsidered. *Experimental Neurology*, 212(1), 226-229. doi:10.1016/j.expneurol.2008.04.001
- Braak, H., Del Tredici, K., Rub, U., de Vos, R. A., Jansen Steur, E. N., & Braak, E. (2003). Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of Aging*, 24(2), 197-211.
- Braak, H., Ghebremedhin, E., Rub, U., Bratzke, H., & Del Tredici, K. (2004). Stages in the development of Parkinson's disease-related pathology. *Cell Tissue Research*, 318(1), 121-134. doi:10.1007/s00441-004-0956-9
- Braak, H., Rub, U., Jansen Steur, E. N., Del Tredici, K., & de Vos, R. A. (2005). Cognitive status correlates with neuropathologic stage in Parkinson disease. *Neurology*, 64(8), 1404-1410. doi:10.1212/01.WNL.0000158422.41380.82
- Brookshire, R.H., & Nicholas, L.E. (1984). Comprehension of directly and indirectly stated main ideas and details in discourse by brain-damaged and non-brain-damaged listeners. *Brain and Language*, 21(1), 21-36.
- Brookshire, R.H., & Nicholas, L.E (1994a). Test-retest stability of measures of connected speech in aphasia. *Clinical Aphasiology*, 22, 119-133.
- Brookshire, R. H., & Nicholas, L.E (1994b). Speech sample size and test-retest stability of connected speech measures for adults with aphasia. *Journal of Speech and Hearing Research*, 37(2), 399-407.
- Brookshire, R. H., & Nicholas, L.E (1997). *The Discourse Comprehension Test - 2nd ed.* Albuquerque, NM: Pica Programs.
- Burn, D. J., Rowan, E. N., Minett, T., Sanders, J., Myint, P., Richardson, J., . . . O'Brien, J. T. (2003). Extrapyramidal features in Parkinson's disease with and without dementia and dementia with Lewy bodies: A cross-sectional comparative study. *Movement disorders*, 18(8), 884-889.
- Cannizzaro, M. S., & Coelho, C. A. (2013). Analysis of narrative discourse structure as an ecologically relevant measure of executive function in adults. *Journal of Psycholinguistic Research*, 42(6), 527-549.
- Capilouto, G., Wright, H. H., & Wagovich, S. A. (2005). CIU and main event analyses of the structured discourse of older and younger adults. *Journal of Communication Disorders*, 38(6), 431-444. doi:10.1016/j.jcomdis.2005.03.005
- Carbon, M., & Marie, R. M. (2003). Functional imaging of cognition in Parkinson's disease. *Current Opinions in Neurology*, 16(4), 475-480. doi:10.1097/01.wco.0000084225.82329.3c
- Chenery, H. J., Angwin, A. J., & Copland, D. A. (2008). The basal ganglia circuits, dopamine, and ambiguous word processing: a neurobiological account of priming studies in Parkinson's disease. *Journal of the International Neuropsychological Society*, 14(3), 351-364. doi:10.1017/S1355617708080491
- Cherney, L. (1998). Pragmatics and Discourse An Introduction. In L. Cherney, B. Shadden, & C. Coelho (Eds.), *Analyzing Discourse In Communicatively Impaired Adults* (pp. 1-8). Gaithersburg: Aspen Publishers.
- Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290.
- Coelho, C. A. (2007). Management of discourse deficits following traumatic brain injury:

- progress, caveats, and needs. *Seminars in Speech and Language*, 28(2), 122-135. doi:10.1055/s-2007-970570 10.1006/brln.1999.2243
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences* (Second Edition ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1), 155-159.
- Colman, K. S., Koerts, J., van Beilen, M., Leenders, K. L., Post, W. J., & Bastiaanse, R. (2009). The impact of executive functions on verb production in patients with Parkinson's disease. *Cortex*, 45(8), 930-942. doi:10.1016/j.cortex.2008.12.010
- Constantinescu, G., Theodoros, D., Russell, T., Ward, E., Wilson, S., & Wootton, R. (2010). Assessing disordered speech and voice in Parkinson's disease: a telerehabilitation application. *International Journal of Language & Communication Disorders*, 45(6), 630-644.
- Cooper, J., Sagar, H., Jordan, N., Harvey, N., & Sullivan, E. (1991). Cognitive impairment in early, untreated Parkinson's disease and its relationship to motor disability. *Brain*, 114(5), 2095-2122.
- Cooper, P. V. (1990). Discourse production and normal aging: Performance on oral picture description tasks. *Journal of Gerontology*, 45(5), P210-P214.
- IBM Corp. (2013). *IBM SPSS Statistics for Windows* (Version 22.0). Armonk, NY: IBM Corp.
- Cotelli, M., Borroni, B., Manenti, R., Zanetti, M., Arevalo, A., Cappa, S. F., & Padovani, A. (2007). Action and object naming in Parkinson's disease without dementia. *European Journal of Neurology*, 14(6), 632-637. doi:10.1111/j.1468-1331.2007.01797.x
- Crescentini, C., Mondolo, F., Biasutti, E., & Shallice, T. (2008). Supervisory and routine processes in noun and verb generation in nondemented patients with Parkinson's disease. *Neuropsychologia*, 46(2), 434-447. doi:10.1016/j.neuropsychologia.2007.08.021
- Crosson, B. (1985). Subcortical functions in language: a working model. *Brain and Language*, 25(2), 257-292.
- Crosson, B. (2012) Thalamic mechanisms in language: A reconsideration based on recent findings and concepts. *Brain and Language*, 126(1), 73-88. doi:10.1016/j.bandl.2012.06.011
- Crucian, G. P., Huang, L., Barrett, A. M., Schwartz, R. L., Cibula, J. E., Anderson, J. M., . . . Heilman, K. M. (2001). Emotional conversations in Parkinson's disease. *Neurology*, 56(2), 159-165.
- Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor Speech Disorders* (Third Edition). Philadelphia, PA: W.B. Saunders.
- Davis, B., & MacLagan, M. (2010). Pauses, fillers, placeholders and formulaicity in Alzheimer's discourse. In N. Amirdze, B. Davis, & M. MacLagan (Eds.), *Fillers, Pauses and Placeholders* (pp. 189-216). John Benjamins Publishing.
- Del Tredici, K., Rub, U., De Vos, R. A., Bohl, J. R., & Braak, H. (2002). Where does parkinson disease pathology begin in the brain? *Journal of Neuropathology and Experimental Neurology*, 61(5), 413-426.
- Doyle, P. J., Goda, A. J., & Spencer, K. A. (1995). The communicative informativeness and efficiency of connected discourse by adults with aphasia under structured and conversational sampling conditions. *American Journal of Speech-Language*

- Pathology*, 4(4), 130-134.
- Doyle, P. J., McNeil, M., Park, G., Goda, A., Rubenstein, E., Spencer, K., . . . Szwarc, L. (2000). Linguistic validation of four parallel forms of a story retelling procedure. *Aphasiology*, 14(5-6), 537-549. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/026870300401306>
- Doyle, P. J., McNeil, M. R., Spencer, K. A., Goda, A. J., Cottrell, K., & Lustig, A. P. (1998). The effects of concurrent picture presentations on retelling of orally presented stories by adults with aphasia. *Aphasiology*, 12(7-8), 561-574. Retrieved from <http://aphasiology.pitt.edu/archive/00001124/01/27-06.pdf>
- Duffau, H. (2008). The anatomo-functional connectivity of language revisited. New insights provided by electrostimulation and tractography. *Neuropsychologia*, 46(4), 927-934. doi:10.1016/j.neuropsychologia.2007.10.025
- Duffau, H., Capelle, L., Denvil, D., Gatignol, P., Sichez, N., Lopes, M., . . . Van Effenterre, R. (2003). The role of dominant premotor cortex in language: a study using intraoperative functional mapping in awake patients. *Neuroimage*, 20(4), 1903-1914.
- Duffau, H., Gatignol, P., Mandonnet, E., Peruzzi, P., Tzourio-Mazoyer, N., & Capelle, L. (2005). New insights into the anatomo-functional connectivity of the semantic system: a study using cortico-subcortical electrostimulations. *Brain*, 128(Pt 4), 797-810. doi:10.1093/brain/awh423
- Duong, A., Tardif, A., & Ska, B. (2003). Discourse about discourse: what is it and how does it progress in Alzheimer's disease? *Brain and Cognition*, 53(2), 177-180.
- Elgh, E., Domellof, M., Linder, J., Edstrom, M., Stenlund, H., & Forsgren, L. (2009). Cognitive function in early Parkinson's disease: a population-based study. *European Journal of Neurology*, 16(12), 1278-1284. doi:10.1111/j.1468-1331.2009.02707.x
- Ellmore, T. M., Beauchamp, M. S., O'Neill, T. J., Dreyer, S., & Tandon, N. (2009). Relationships between essential cortical language sites and subcortical pathways. *Journal Neurosurgery*, 111(4), 755-766. doi:10.3171/2009.3.JNS081427
- Emre, M., Aarsland, D., Brown, R., Burn, D. J., Duyckaerts, C., Mizuno, Y., . . . Dubois, B. (2007). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Movement Disorders*, 22(12), 1689-707; quiz 1837. doi:10.1002/mds.21507
- Engle, R. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11(1), 19-23.
- Esselink, R. A. J., de Bie, R. M. A., Schuurman, P. R., & Speelman, J. D. (2007). Parkinson's disease-related Stereotactic Functional Neurosurgery. In E. C. Wolters, T. van Laar, & H. W. Berendse (Eds.), *Parkinsonism and Related Disorders* (pp. 241-249).
- Fahn, S. R. L. E., Elton, R., & Committee, U. P. D. R. S. D. (1987). Unified Parkinson's disease rating scale. *Recent developments in Parkinson's disease*, 2, 153-163.
- Farag, C., Troiani, V., Bonner, M., Powers, C., Avants, B., Gee, J., & Grossman, M. (2010). Hierarchical organization of scripts: converging evidence from fMRI and frontotemporal degeneration. *Cerebral Cortex*, 20(10), 2453-2463. doi:10.1093/cercor/bhp313
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible

- statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods*, 39(2), 175-191.
- Fergadiotis, G., & Wright, H. H. (2011). Lexical diversity for adults with and without aphasia across discourse elicitation tasks. *Aphasiology*, 25(11), 1414-1430. doi:10.1080/02687038.2011.603898
- Fergadiotis, G., Wright, H. H., & Capilouto, G. J. (2011). Productive vocabulary across discourse types. *Aphasiology*, 25(10), 1261-1278. doi:10.1080/02687038.2011.606974
- Fergadiotis, G., Wright, H. H., & West, T. M. (2013). Measuring lexical diversity in narrative discourse of people with aphasia. *American Journal of Speech Language Pathology*, 22(2), 397-408. doi:10.1044/1058-0360(2013/12-0083)
- Fernandino, L., Conant, L. L., Binder, J. R., Blindauer, K., Hiner, B., Spangler, K., & Desai, R. H. (2012). Parkinson's disease disrupts both automatic and controlled processing of action verbs. *Brain and Language*.
- Fleming, V. B., & Harris, J. L. (2008). Complex discourse production in mild cognitive impairment: Detecting subtle changes. *Aphasiology*, 22(7-8), 729-740.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189-198.
- Gago, M. F., Garrett, M. C., Fonseca, M. R., Rosas, M. J., Simoes, M. F., Vieira, S., & Botelho, F. (2009). How do cognitive and axial motor signs correlate in Parkinson's disease? A 6-year prospective study. *Journal of Neurology*, 256(10), 1655-1662. doi:10.1007/s00415-009-5174-7
- Gelb, D. J., Oliver, E., & Gilman, S. (1999). Diagnostic criteria for Parkinson disease. *Archives of Neurology*, 56(1), 33-39.
- German, D. J. (1989). *Test of adolescent/adult word finding*. Austin, TX: Pro-Ed.
- Glosser, G., & Deser, T. (1992). A comparison of changes in macrolinguistic and microlinguistic aspects of discourse production in normal aging. *Journal of Gerontology*, 47(4), P266-P272.
- Godbout, L., & Doyon, J. (2000). Defective representation of knowledge in Parkinson's disease: evidence from a script-production task. *Brain and Cognition*, 44(3), 490-510. doi:10.1006/breg.2000.1213
- Goetz, C. G., Poewe, W., Rascol, O., Sampaio, C., Stebbins, G. T., Counsell, C., . . . Seidl, L. (2004). Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. *Movement Disorders*, 19(9), 1020-1028. doi:10.1002/mds.20213
- Golden, C., & Freshwater, S. (2002). *The Stroop Color and Word Test: A Manual for Clinical and Experimental Uses*. Wood Dale, IL USA: Stoelting Co.
- Graves, R. E., Bezeau, S. C., Fogarty, J., & Blair, R. (2004). Boston naming test short forms: a comparison of previous forms with new item response theory based forms. *Journal of Clinical and Experimental Neuropsychology*, 26(7), 891-902. doi:10.1080/13803390490510716
- Green, J., McDonald, W. M., Vitek, J. L., Evatt, M., Freeman, A., Haber, M., . . . DeLong, M. R. (2002). Cognitive impairments in advanced PD without dementia. *Neurology*, 59(9), 1320-1324.
- Griffiths, S., Barnes, R., Britten, N., & Wilkinson, R. (2012). Potential causes and

- consequences of overlap in talk between speakers with Parkinson's disease and their familiar conversation partners. *Seminars in Speech and Language*, 33(1), 27-43. doi:10.1055/s-0031-1301161
- Groenewegen, H. J. (2003). The basal ganglia and motor control. *Neural Plasticity*, 10(1-2), 107-120. doi:10.1155/NP.2003.107
- Grossman, M. (1999). Sentence processing in Parkinson's disease. *Brain and Cognition*, 40(2), 387-413. doi:10.1006/breg.1999.1087
- Grossman, M., Carvell, S., Gollomp, S., Stern, M. B., Vernon, G., & Hurtig, H. I. (1991). Sentence comprehension and praxis deficits in Parkinson's disease. *Neurology*, 41(10), 1620-1626.
- Grossman, M., Carvell, S., & Peltzer, L. (1993). The sum and substance of it: the appreciation of mass and count quantifiers in Parkinson's disease. *Brain and Language*, 44(4), 351-384. doi:10.1006/brln.1993.1022
- Grossman, M., Carvell, S., Stern, M. B., Gollomp, S., & Hurtig, H. I. (1992a). Sentence comprehension in Parkinson's disease: the role of attention and memory. *Brain and Language*, 42(4), 347-384.
- Grossman, M., Cooke, A., DeVita, C., Lee, C., Alsop, D., Detre, J., . . . Hurtig, H. I. (2003). Grammatical and resource components of sentence processing in Parkinson's disease: an fMRI study. *Neurology*, 60(5), 775-781.
- Grossman, M., Crino, P., Reivich, M., Stern, M. B., & Hurtig, H. I. (1992b). Attention and sentence processing deficits in Parkinson's disease: the role of anterior cingulate cortex. *Cerebral Cortex*, 2(6), 513-525.
- Grossman, M., Gross, R. G., Moore, P., Dreyfuss, M., McMillan, C. T., Cook, P. A., . . . Siderowf, A. (2012). Difficulty processing temporary syntactic ambiguities in Lewy body spectrum disorder. *Brain and Language*, 120(1), 52-60. doi:10.1016/j.bandl.2011.08.007
- Grossman, M., Kalmanson, J., Bernhardt, N., Morris, J., Stern, M. B., & Hurtig, H. I. (2000). Cognitive resource limitations during sentence comprehension in Parkinson's disease. *Brain and Language*, 73(1), 1-16. doi:10.1006/brln.2000.2290
- Hakstian, A. R., Roed, J. C., & Lind, J. C. (1979). Two-sample T-2 procedure and the assumption of homogeneous covariance matrices. *Psychological Bulletin*, 86(6), 1255.
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in quantitative methods for psychology*, 8(1), 23.
- Halpern, C. H., Rick, J. H., Danish, S. F., Grossman, M., & Baltuch, G. H. (2009). Cognition following bilateral deep brain stimulation surgery of the subthalamic nucleus for Parkinson's disease. *International Journal Geriatric Psychiatry*, 24(5), 443-451. doi:10.1002/gps.2149
- Hare, M., Jones, M., Thomson, C., Kelly, S., & McRae, K. (2009). Activating event knowledge. *Cognition*, 111(2), 151-167. doi:10.1016/j.cognition.2009.01.009
- Heap, H. (1978). *Information retrieval: Computational and theoretical aspects*. New York, NY: Academic Press, Inc.
- Helm-Estabrooks, N. (1992). *Aphasia Diagnostic Profiles*. Chicago, IL: Riverside.
- Henry, J. D., & Crawford, J. R. (2004). Verbal fluency deficits in Parkinson's disease: a meta-analysis. *Journal of the International Neuropsychological Society*, 10(4), 608-622. doi:10.1017/S1355617704104141

- Herrera, E., & Cueto, F. (2012). Action naming in Parkinson's disease patients on/off dopamine. *Neuroscience Letters*, 513(2), 219-222. doi:10.1016/j.neulet.2012.02.045
- Herrera, E., Rodriguez-Ferreiro, J., & Cueto, F. (2012). The effect of motion content in action naming by Parkinson's disease patients. *Cortex*, 48(7), 900-904. doi:10.1016/j.cortex.2010.12.007
- Herrero, M. T., Barcia, C., & Navarro, J. M. (2002). Functional anatomy of thalamus and basal ganglia. *Child's Nervous System*, 18(8), 386-404. doi:10.1007/s00381-002-0604-1
- Hochstadt, J. (2009). Set-shifting and the on-line processing of relative clauses in Parkinson's disease: results from a novel eye-tracking method. *Cortex*, 45(8), 991-1011. doi:10.1016/j.cortex.2009.03.010
- Hochstadt, J., Nakano, H., Lieberman, P., & Friedman, J. (2006). The roles of sequencing and verbal working memory in sentence comprehension deficits in Parkinson's disease. *Brain and Language*, 97(3), 243-257. doi:10.1016/j.bandl.2005.10.011
- Holdorff, B. (2002). Friedrich Heinrich Lewy (1885-1950) and his work. *Journal of the Histories of the Neurosciences*, 11(1), 19-28. doi:10.1076/jhin.11.1.19.9106
- Holtgraves, T., & McNamara, P. (2010). Pragmatic comprehension deficit in Parkinson's disease. *Journal of Clinical Experimental Neuropsychology*, 32(4), 388-397. doi:10.1080/13803390903130729
- Holtgraves, T., McNamara, P., Cappaert, K., & Durso, R. (2010). Linguistic correlates of asymmetric motor symptom severity in Parkinson's disease. *Brain and Cognition*, 72(2), 189-196.
- Hornykiewicz, O. (2001). Chemical neuroanatomy of the basal ganglia--normal and in Parkinson's disease. *Journal of Chemical Neuroanatomy*, 22(1-2), 3-12.
- Howard, D., & Patterson, K. (1992). *The Pyramids and Palm Trees Test*. London, UK: Pearson Assessment.
- Huber, J. E., & Darling, M. (2011). Effect of Parkinson's disease on the production of structured and unstructured speaking tasks: respiratory physiologic and linguistic considerations. *Journal Speech and Language and Hearing Research*, 54(1), 33-46. doi:10.1044/1092-4388(2010/09-0184)
- Hughes, A. J., Daniel, S. E., Kilford, L., & Lees, A. J. (1992). Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *Journal of Neurology, Neurosurgery and Psychiatry*, 55(3), 181-184.
- Hula, W., McNeil, M., Doyle, P., Rubinsky, H., & Fossett, T. (2003). The inter-rater reliability of the story retell procedure. *Aphasiology*, 17(5), 523-528.
- Hummel, T. J., & Sligo, J. R. (1971). Empirical comparison of univariate and multivariate analysis of variance procedures. *Psychological Bulletin*, 76(1), 49.
- Illes, J. (1989). Neurolinguistic features of spontaneous language production dissociate three forms of neurodegenerative disease: Alzheimer's, Huntington's, and Parkinson's. *Brain and Language*, 37(4), 628-642.
- Illes, J., Metter, E. J., Hanson, W. R., & Iritani, S. (1988). Language production in Parkinson's disease: acoustic and linguistic considerations. *Brain and Language*, 33(1), 146-160.
- Ivory, S. J., Knight, R. G., Longmore, B. E., & Caradoc-Davies, T. (1999). Verbal memory in non-demented patients with idiopathic Parkinson's disease. *Neuropsychologia*, 37(7), 817-828.

- Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. *Journal of Neurology, Neurosurgery, and Psychiatry*, 79(4), 368-376.
doi:10.1136/jnnp.2007.131045
- Jaywant, A., & Pell, M. D. (2010). Listener impressions of speakers with Parkinson's disease. *Journal of the International Neuropsychological Society*, 16(1), 49-57.
doi:10.1017/S1355617709990919
- Jensen, A. M., Chenery, H. J., & Copland, D. A. (2006). A comparison of picture description abilities in individuals with vascular subcortical lesions and Huntington's Disease. *Journal of Communication Disorders*, 39(1), 62-77.
doi:10.1016/j.jcomdis.2005.07.001
- Jokel, R., De Nil, L., & Sharpe, K. (2007). Speech disfluencies in adults with neurogenic stuttering associated with stroke and traumatic brain injury. *Journal of Medical Speech-Language Pathology*, 15(3), 243-261.
- Jurica, P. J., Leitten, C. L., & Mattis, S. (2001). *DRS-2: Dementia Rating Scale-2 Professional Manual*. Odessa, Florida: Psychological Assessment Resources Eds.
- Kalbe, E., Calabrese, P., Kohn, N., Hilker, R., Riedel, O., Wittchen, H.-U., . . . Kessler, J. (2008). Screening for cognitive deficits in Parkinson's disease with the Parkinson neuropsychometric dementia assessment (PANDA) instrument. *Parkinsonism & Related Disorders*, 14(2), 93-101.
- Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston naming test* (Second Edition). Austin, TX: Pro-ed.
- Kempler, D., & Van Lancker, D. V. (2002). Effect of speech task on intelligibility in dysarthria: a case study of Parkinson's Disease. *Brain and Language*, 80(3), 449-464.
- Khan, A., & Rayner, G. D. (2003). Robustness to non-normality of common tests for the many-sample location problem. *Journal of Applied Mathematics & Decision Sciences*, 7(4), 187-206.
- Kielar, A., Milman, L., Bonakdarpour, B., & Thompson, C. K. (2011). Neural correlates of covert and overt production of tense and agreement morphology: Evidence from fMRI. *Journal of Neurolinguistics*, 24(2), 183-201.
- Koerts, J., Meijer, H. A., Colman, K. S., Tucha, L., Lange, K. W., & Tucha, O. (2012). What is measured with verbal fluency tests in Parkinson's disease patients at different stages of the disease? *Journal of Neural Transmission*, 120(3), 403-411.
doi:10.1007/s00702-012-0885-9
- Kudlicka, A., Clare, L., & Hindle, J. V. (2011). Executive functions in Parkinson's disease: Systematic review and meta-analysis. *Movement Disorders*, 26(13), 2305-2315.
- Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. *The Neurocognition of Language*, 83-122.
- Lewis, F. M., Lapointe, L. L., Murdoch, B. E., & Chenery, H. J. (1998). Language impairment in Parkinson's disease. *Aphasiology*, 12(3), 193-206.
- Light, L. & Anderson, P. (1985). Working memory capacity, age, and memory for discourse. *Journal of Gerontology*, 40(6), 737-747.
- Lippa, C. F., Duda, J. E., Grossman, M., Hurtig, H. I., Aarsland, D., Boeve, B. F., . . . Emre, M. (2007). DLB and PDD boundary issues Diagnosis, treatment, molecular pathology, and biomarkers. *Neurology*, 68(11), 812-819.

- Litvan, I., Goldman, J. G., Troster, A. I., Schmand, B. A., Weintraub, D., Petersen, R. C., . . . Emre, M. (2012). Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines. *Movement Disorders*, 27(3), 349-356. doi:10.1002/mds.24893
- Llebaria, G., Pagonabarraga, J., Kulisevsky, J., Garcia-Sanchez, C., Pascual-Sedano, B., Gironell, A., & Martinez-Corral, M. (2008). Cut-off score of the Mattis Dementia Rating Scale for screening dementia in Parkinson's disease. *Movement Disorders*, 23(11), 1546-1550. doi:10.1002/mds.22173
- Locascio, J. J., Corkin, S., & Growdon, J. H. (2003). Relation between clinical characteristics of Parkinson's disease and cognitive decline. *Journal of Clinical Experimental Neuropsychology*, 25(1), 94-109. doi:10.1076/jcen.25.1.94.13624
- Lucas, J. A., Ivnik, R. J., Smith, G. E., Bohac, D. L., Tangalos, E. G., Kokmen, E., . . . Petersen, R. C. (1998). Normative data for the Mattis dementia rating scale. *Journal of Clinical and Experimental Neuropsychology*, 20(4), 536-547.
- MacDonald, P., MacDonald, A., Seergobin, K., Tamjeedi, R., Ganjavi, H., Provost, J., & Monchi, O. (2011). The effect of dopamine therapy on ventral and dorsal striatum-mediated cognition in Parkinson's disease: Support from functional MRI. *Brain*, 134, 1447-1463. doi: 10.1093/brain/awr075
- Mackenzie, C. (2000). Adult spoken discourse: The influences of age and education. *International Journal of Language & Communication Disorders*, 35(2), 269-285.
- MacWhinney, B. (2014). The CHILDES Project Tools for Analyzing Talk - Electronid Edition.
- Mardia, K. V. (1971). The effect of nonnormality on some multivariate tests and robustness to nonnormality in the linear model. *Biometrika*, 58(1), 105-121.
- Marini, A., Boewe, A., Caltagirone, C., & Carlomagno, S. (2005). Age-related differences in the production of textual descriptions. *Journal of Psycholinguistic Research*, 34(5), 439-463. doi: 10.1007/s10936-005-6203-z
- Marras, C., Armstrong, M. J., Meaney, C. A., Fox, S., Rothberg, B., Reginold, W., . . . Duff-Canning, S. (2013). Measuring mild cognitive impairment in patients with Parkinson's disease. *Movement Disorders*, 28(5), 626-633. doi:10.1002/mds.25426
- Marsh, L., McDonald, W. M., Cummings, J., & Ravina, B. (2006). Provisional diagnostic criteria for depression in Parkinson's disease: report of an NINDS/NIMH Work Group. *Movement Disorders*, 21(2), 148-158.
- Martin, W. R., & Wieler, M. (2003). Treatment of Parkinson's disease. *Canadian Journal of Neurological Sciences*, 30 Suppl 1, S27-S33.
- Matison, R., Mayeux, R., Rosen, J., & Fahn, S. (1982). "Tip-of-the-tongue" phenomenon in Parkinson disease. *Neurology*, 32(5), 567-570.
- Mattis, S. (1988). *Dementia Rating Scale (DRS)*. Odessa, FL: Psychological Assessment Resources.
- Mattis, S. (2001). *Mattis Dementia Rating Scale - 2*. Odessa, FL: Psychological Assessment Resources.
- Mayer, J. & Murray, L.L. (2003). Functional measures of naming in aphasia: Word retrieval in confrontation naming versus connected speech. *Aphasiology*, 17(5), 481-497.
- McKeith, I. G., & Mosimann, U. P. (2004). Dementia with Lewy bodies and Parkinson's

- disease. *Parkinsonism & Related Disorders*, 10 Suppl 1, S15-S18.
doi:10.1016/j.parkreldis.2003.12.005
- McKinlay, A., Dalrymple-Alford, J. C., Grace, R. C., & Roger, D. (2009). The effect of attentional set-shifting, working memory, and processing speed on pragmatic language functioning in Parkinson's disease. *European Journal of Cognitive Psychology*, 21(2-3), 330-346.
- McNamara, P., & Durso, R. (2003). Pragmatic communication skills in patients with Parkinson's disease. *Brain and Language*, 84(3), 414-423.
- McNamara, P., Durso, R., Brown, A., & Lynch, A. (2003). Counterfactual cognitive deficit in persons with Parkinson's disease. *Journal Neurology, Neurosurgery, and Psychiatry*, 74(8), 1065-1070.
- McNamara, P., Obler, L. K., Au, R., Durso, R., & Albert, M. L. (1992). Speech monitoring skills in Alzheimer's disease, Parkinson's disease, and normal aging. *Brain and Language*, 42(1), 38-51.
- McNeil, M. R., Doyle, P. J., Fossett, T. R. D., Park, G. H., & Goda, A. J. (2001). Reliability and concurrent validity of the information unit scoring metric for the story retelling procedure. *Aphasiology*, 15(10-11), 991-1006.
- McNeil, M. R., Sung, J. E., Yang, D., Pratt, S. R., Fossett, T. R. D., Doyle, P. J., & Pavelko, S. (2007). Comparing connected language elicitation procedures in persons with aphasia: Concurrent validation of the Story Retell Procedure. *Aphasiology*, 21(6-8), 775-790.
- Merlo, S., & Mansur, L. L. (2004). Descriptive discourse: topic familiarity and disfluencies. *Journal of Communication Disorders*, 37(6), 489-503.
- Middleton, F. A., & Strick, P. L. (2000). Basal ganglia output and cognition: evidence from anatomical, behavioral, and clinical studies. *Brain and Cognition*, 42(2), 183-200. doi:10.1006/brcg.1999.1099
- Miller, N., Noble, E., Jones, D., & Burn, D. (2006). Life with communication changes in Parkinson's disease. *Age and Ageing*, 35(3), 235-239. doi:10.1093/ageing/afj053
- Monetta, L., Cheang, H. S., & Pell, M. D. (2008). Understanding speaker attitudes from prosody by adults with Parkinson's disease. *Journal of Neuropsychology*, 2(Pt 2), 415-430.
- Monetta, L., Grindrod, C. M., & Pell, M. D. (2009). Irony comprehension and theory of mind deficits in patients with Parkinson's disease. *Cortex*, 45(8), 972-981. doi:10.1016/j.cortex.2009.02.021
- Monetta, L., & Pell, M. D. (2007). Effects of verbal working memory deficits on metaphor comprehension in patients with Parkinson's disease. *Brain and Language*, 101(1), 80-89. doi:10.1016/j.bandl.2006.06.007
- Müller, J., Wenning, G. K., Verny, M., McKee, A., Chaudhuri, K. R., Jellinger, K., . . . Litvan, I. (2001). Progression of dysarthria and dysphagia in postmortem-confirmed parkinsonian disorders. *Archives of Neurology*, 58(2), 259-264.
- Murray, L. L. (2000). Spoken language production in Huntington's and Parkinson's diseases. *Journal of Speech Language and Hearing Research*, 43(6), 1350-1366.
- Murray, L.L. (2010). Distinguishing clinical depression from early Alzheimer's disease in elderly people: Can narrative analysis help. *Aphasiology*, 24(6-8), 928-939.
- Murray, L. L., & Stout, J. C. (1999). Discourse comprehension in Huntington's and Parkinson's diseases. *American Journal of Speech-Language Pathology*, 8(2), 137.

- Murray, L. L., & Lenz, L. P. (2001). Productive syntax abilities in Huntington's and Parkinson's diseases. *Brain and Cognition*, 46(1-2), 213-219.
- Muslimovic, D., Post, B., Speelman, J. D., & Schmand, B. (2005). Cognitive profile of patients with newly diagnosed Parkinson disease. *Neurology*, 65(8), 1239-1245. doi:10.1212/01.wnl.0000180516.69442.95
- Nadeau, S. E. (2008). Subcortical Language Mechanisms. In B. Stemmer & H. Whitaker (Eds.), *Handbook of the Neuroscience of Language* (pp. 329-338). London, UK: Academic Press Elsevier.
- Natsopoulos, D., Grouios, G., Bostantzopoulou, S., Mentenopoulos, G., Katsarou, Z., & Logothetis, J. (1993). Algorithmic and heuristic strategies in comprehension of complement clauses by patients with Parkinson's disease. *Neuropsychologia*, 31(9), 951-964.
- Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech and Hearing Research*, 36(2), 338-350.
- Nicholas, L. E., & Brookshire, R. H. (1995). Presence, completeness, and accuracy of main concepts in the connected speech of non-brain-damaged adults and adults with aphasia. *Journal of Speech and Hearing Research*, 38(1), 145-156.
- North, A. J., Ulatowska, H. K., Macaluso-Haynes, S., & Bell, H. (1986). Discourse performance in older adults. *International Journal of Aging and Human Development*, 23(4), 267-283.
- Obeso, I., Casabona, E., Bringas, M. L., Alvarez, L., & Jahanshahi, M. (2012). Semantic and phonemic verbal fluency in Parkinson's disease: Influence of clinical and demographic variables. *Behavioural Neurology*, 25(2), 111-118.
- Obler, L. K., Au, R., Kugler, J., Melvold, J., Tocco, M., & Albert, M. L. (1994). Intersubject variability in adult normal discourse. In *Discourse analysis and applications: Studies in adult clinical populations* (pp. 15-27). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Okazaki, H., Lipkin, L. E., Aronson, S. M. (1961). Diffuse intracytoplasmic ganglionic inclusions (Lewy type) associated with progressive dementia and quadriplegia in flexion. *Journal of Neuropathology and Experimental Neurology*, 20, 237-244.
- Olanow, C. W., & Tatton, W. G. (1999). Etiology and pathogenesis of Parkinson's disease. *Annual Review of Neuroscience*, 22, 123-144. doi:10.1146/annurev.neuro.22.1.123
- Orange, J. B., & Kertesz, A. (2000). Discourse analyses and dementia. *Brain and Language*, 71(1), 172-174. doi:10.1006/brln.1999.2243
- Peach, R. (2013). The cognitive basis for sentence planning difficulties in discourse after traumatic brain injury. *American Journal of Speech-Language Pathology*, 22 S285-S297.
- Pagonabarraga, J., & Kulisevsky, J. (2012). Cognitive impairment and dementia in Parkinson's disease. *Neurobiology of Disease*, 46(3), 590-596. doi:10.1016/j.nbd.2012.03.029
- Pagonabarraga, J., Kulisevsky, J., Llebaria, G., Garcia-Sanchez, C., Pascual-Sedano, B., & Gironell, A. (2008). Parkinson's disease-cognitive rating scale: a new cognitive scale specific for Parkinson's disease. *Movement Disorders*, 23(7), 998-1005. doi:10.1002/mds.22007

- Parkinson Society Canada (2003). Parkinson's Disease: Social and Economic Impact.
- Parkinson, J. (2002). An essay on the shaking palsy. 1817. *Journal of Neuropsychiatry and Clinical Neuroscience*, 14(2), 223-36.
- Parsons, T. D., Rizzo, A. R., Zaag, C. V. D., McGee, J. S., & Buckwalter, J. G. (2005). Gender differences and cognition among older adults. *Aging, Neuropsychology, and Cognition*, 12(1), 78-88.
- Pavon, J. M., Whitson, H. E., & Okun, M. S. (2010). Parkinson's disease in women: a call for improved clinical studies and for comparative effectiveness research. *Maturitas*, 65(4), 352-358.
- Pell, M. D., Cheang, H. S., & Leonard, C. L. (2006). The impact of Parkinson's disease on vocal-prosodic communication from the perspective of listeners. *Brain and language*, 97(2), 123-134.
- Peran, P., Cardebat, D., Cherubini, A., Piras, F., Luccichenti, G., Peppe, A., . . . Sabatini, U. (2009). Object naming and action-verb generation in Parkinson's disease: a fMRI study. *Cortex*, 45(8), 960-971. doi:10.1016/j.cortex.2009.02.019
- Peran, P., Demonet, J. F., Cherubini, A., Cardebat, D., Caltagirone, C., & Sabatini, U. (2010). Mental representations of action: the neural correlates of the verbal and motor components. *Brain Research*, 1328, 89-103. doi:10.1016/j.brainres.2010.02.082
- Pereira, J. B., Junque, C., Marti, M. J., Ramirez-Ruiz, B., Bartres-Faz, D., & Tolosa, E. (2009). Structural brain correlates of verbal fluency in Parkinson's disease. *Neuroreport*, 20(8), 741-744. doi:10.1097/WNR.0b013e328329370b
- Péron, J., Vicente, S., Leray, E., Drapier, S., Drapier, D., Cohen, R., . . . Sauleau, P. (2009). Are dopaminergic pathways involved in theory of mind? A study in Parkinson's disease. *Neuropsychologia*, 47(2), 406-414.
- Piatt, A. L., Fields, J. A., Paolo, A. M., Koller, W. C., & Tröster, A. I. (1999a). Lexical, semantic, and action verbal fluency in Parkinson's disease with and without dementia. *Journal of Clinical and Experimental Neuropsychology*, 21(4), 435-443.
- Piatt, A. L., Fields, J. A., Paolo, A. M., & Tröster, A. I. (1999b). Action (verb naming) fluency as an executive function measure: convergent and divergent evidence of validity. *Neuropsychologia*, 37(13), 1499-1503.
- Pinto, S., Ozsancak, C., Tripoliti, E., Thobois, S., Limousin-Dowsey, P., & Auzou, P. (2004). Treatments for dysarthria in Parkinson's disease. *Lancet Neurology*, 3(9), 547-556. doi:10.1016/S1474-4422(04)00854-3
- Pitcairn, T. K., Clemie, S., Gray, J. M., & Pentland, B. (1990). Impressions of parkinsonian patients from their recorded voices. *British Journal of Disorders of Communication*, 25(1), 85-92.
- Poletti, M., & Bonuccelli, U. (2013). Acute and chronic cognitive effects of levodopa and dopamine agonists on patients with Parkinson's disease: a review. *Therapeutic Advances in Psychopharmacology*, 3(2), 101-113.
- Post, B., Merkus, M. P., de Bie, R., de Haan, R. J., & Speelman, J. D. (2005). Unified Parkinson's disease rating scale motor examination: are ratings of nurses, residents in neurology, and movement disorders specialists interchangeable? *Movement Disorders*, 20(12), 1577-1584.
- Postuma, R. B., Gagnon, J. F., & Montplaisir, J. (2010). Clinical prediction of Parkinson's disease: planning for the age of neuroprotection. *Journal of Neurology*,

- Neurosurgery & Psychiatry*, 81(9), 1008-1013.
- Pulvermuller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, 6(7), 576-582. doi:10.1038/nrn1706
- Pulvermuller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience*, 17(6), 884-892. doi:10.1162/0898929054021111
- Ramig, L. O., Fox, C., & Sapir, S. (2004). *Parkinson's disease: speech and voice disorders and their treatment with the Lee Silverman Voice Treatment*. Proceedings from Seminars in Speech and Language.
- Rankin, G., & Stokes, M. (1998). Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clinical Rehabilitation*, 12(3), 187-199.
- Reitan, R. M., & Wolfson, D. (1985). *The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation*. Tuscon, AZ: Neuropsychology Press.
- Roberts-South, A., Findlater, K., Strong, M. J., & Orange, J. B. (2012). Longitudinal changes in discourse production in amyotrophic lateral sclerosis. *Seminars in Speech and Language*, 33(1), 79-94. doi:10.1055/s-0031-1301165
- Roberts, A., Finger, E., Jenkins, M., Eaton, B., & Orange, J. B. (2014). *Communication in Parkinson disease: A family perspective*. Proceedings from SAC Annual Conference, Ottawa, ON Canada.
- Roberts, A., & Orange, J. B. (2013). Discourse in Lewy Body Spectrum Disorder. In B. Davis & J. Guendouzi (Eds.), *Pragmatics in Dementia Discourse* (pp. 147-204). Newcastle upon Tyne: Cambridge Scholars Publishing.
- Roberts, P. M., Meltzer, A., & Wilding, J. (2009). Disfluencies in non-stuttering adults across sample lengths and topics. *Journal of Communication Disorders*, 42(6), 414-427. R
- Rodriguez-Ferreiro, J., Cuetos, F., Herrera, E., Menendez, M., & Ribacoba, R. (2010). Cognitive impairment in Parkinson's disease without dementia. *Movement Disorders*, 25(13), 2136-2141. doi:10.1002/mds.23239
- Rodriguez-Ferreiro, J., Menendez, M., Ribacoba, R., & Cuetos, F. (2009). Action naming is impaired in Parkinson disease patients. *Neuropsychologia*, 47(14), 3271-3274. doi:10.1016/j.neuropsychologia.2009.07.007
- Rousseaux, M., Seve, A., Vallet, M., Pasquier, F., & Mackowiak-Cordoliani, M. A. (2010). An analysis of communication in conversation in patients with dementia. *Neuropsychologia*, 48(13), 3884-3890. doi:10.1016/j.neuropsychologia.2010.09.026
- Sapir, S., Pawlas, A., Ramig, L., Countryman, S., O'Brien, C., Hoehn, M., & Thompson, L. (2001). Voice and speech abnormalities in Parkinson disease: Relation to severity of motor impairment, duration of disease, medication, depression, gender, and age. *Journal of Medical Speech-Language Pathology*, 9(4), 213-226.
- SALT Software, L. L. C. Systematic Analysis of Language Transcription (SALT): On-line resources. Retrieved from <http://www.saltsoftware.com/resources/trainads/#>
- Scarborough, H. S., Rescorla, L., Tager-Flusberg, H., Fowler, A. E., & Sudhalter, V. (1991). The relation of utterance length to grammatical complexity in normal and language-disordered groups. *Applied Psycholinguistics*, 12(1), 23-45.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art.

- Psychological methods*, 7(2), 147.
- Schapira, A. H. (2007). Treatment options in the modern management of Parkinson disease. *Archives of Neurology*, 64(8), 1083-1088. doi:10.1001/archneur.64.8.1083
- Schrag, A., Barone, P., Brown, R. G., Leentjens, A. F. G., McDonald, W. M., Starkstein, S., . . . Sampaio, C. (2007). Depression rating scales in Parkinson's disease: critique and recommendations. *Movement Disorders*, 22(8), 1077-1092.
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge university press.
- Shadden, B. (1998a). Obtaining the Discourse Sample. In L. Cherney, B. Shadden, & C. Coelho (Eds.), *Analyzing Discourse in Communicatively Impaired Adults* (pp. 9-34). Gaithersburg: Aspen Publishers.
- Shadden, B. (1998b). Information Analyses. In L. Cherney, B. Shadden, & C. Coelho (Eds.), *Analyzing Discourse in Communicatively impaired Adults* (pp. 85-114). Gaithersburg: Aspen Publishers.
- Shadden, B. B., Burnette, R. B., Eikenberry, B. R., & DiBrezzo, R. (1991). All discourse tasks are not created equal. *Aphasiology*, 20, 327-341.
- Shulman, L. M., Gruber-Baldini, A. L., Anderson, K. E., Fishman, P. S., Reich, S. G., & Weiner, W. J. (2010). The clinically important difference on the unified Parkinson's disease rating scale. *Archives of Neurology*, 67(1), 64-70.
- Shulman, L. M., Taback, R. L., Rabinstein, A. A., & Weiner, W. J. (2002). Non-recognition of depression and other non-motor symptoms in Parkinson's disease. *Parkinsonism & Related Disorders*, 8(3), 193-197.
- Siderowf, A., McDermott, M., Kieburtz, K., Blindauer, K., Plumb, S., & Shoulson, I. (2002). Test-retest reliability of the unified Parkinson's disease rating scale in patients with early Parkinson's disease: results from a multicenter clinical trial. *Movement Disorders*, 17(4), 758-763.
- Signorini, M., & Volpato, C. (2006). Action fluency in Parkinson's disease: a follow-up study. *Movement Disorders*, 21(4), 467-472. doi:10.1002/mds.20718
- Skeel, R. L., Crosson, B., Nadeau, S. E., Algina, J., Bauer, R. M., & Fennell, E. B. (2001). Basal ganglia dysfunction, working memory, and sentence comprehension in patients with Parkinson's disease. *Neuropsychologia*, 39(9), 962-971.
- Solla, P., Cannas, A., Ibba, F. C., Loi, F., Corona, M., Orofino, G., . . . Marrosu, F. (2012). Gender differences in motor and non-motor symptoms among Sardinian patients with Parkinson's disease. *Journal of Neurological Sciences*, 323 (1), 33-39.
- Sollinger, A. B., Goldstein, F. C., Lah, J. J., Levey, A. I., & Factor, S. A. (2010). Mild cognitive impairment in Parkinson's disease: subtypes and motor characteristics. *Parkinsonism & Related Disorders*, 16(3), 177-180. doi:10.1016/j.parkreldis.2009.11.002
- Song, Y., Gu, Z., An, J., & Chan, P. (2014). Gender differences on motor and non-motor symptoms of de novo patients with early Parkinson's disease. *Neurological Sciences*. Published ahead of print. Retrieved from: <http://link.springer.com/article/10.1007/s10072-014-1879-1/fulltext.html>
- Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms and commentary* (2nd ed.). New York, NY USA: Oxford University Press.
- Stebbins, G., Goetz, C., Burn, D., Jankovic, J., Khoo, T., & Tilley, B. (2013). How to

- identify tremor dominant and postural instability/gait difficulty groups with the Movement Disorder Society Unified Parkinson's Disease Rating Scale: Comparison with the Unified Parkinson's Disease Rating Scale. *Movement Disorders*, 28(5), 688-670.
- Strauss, E., Sherman, E., & Spreen, O. (2006). General Cognitive Functioning, Neuropsychological Batteries, and Assessment of Premorbid Intelligence: Dementia Rating Scale - 2. In E. Strauss, E. Sherman, & O. Spreen (Eds.), *A Compendium of Neuropsychological Tests* (Third ed., pp. 144-158). New York, New York: Oxford University Press.
- Svenningsson, P., Westman, E., Ballard, C., & Aarsland, D. (2012). Cognitive impairment in patients with Parkinson's disease: diagnosis, biomarkers, and treatment. *Lancet Neurology*, 11(8), 697-707. doi:10.1016/S1474-4422(12)70152-7
- Tabachnick, B. G., & Fidell, L. S. (2006). *Using multivariate statistics* (Fifth Edition ed.). Boston: Allyn and Bacon Boston.
- Teichmann, M., Darcy, I., Bachoud-Levi, A. C., & Dupoux, E. (2009). The role of the striatum in phonological processing: evidence from early stages of Huntington's disease. *Cortex*, 45(7), 839-849. doi:10.1016/j.cortex.2008.12.005
- Templin, M. (1957). *Certain Language Skills in Children: Their Development and Relationships*. Minneapolis: University of Minnesota Press.
- Thenganatt, M. A., & Jankovic, J. (2014). Parkinson disease subtypes. *JAMA Neurology*, 71(4), 499-504.
- Thompson, C. K. (2011). Northwestern assessment of verbs and sentences. In. Chicago, Illinois: Northwestern University.
- Tomlinson, C. L., Stowe, R., Patel, S., Rick, C., Gray, R., & Clarke, C. E. (2010). Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Movement Disorders*, 25(15), 2649-2653. doi:10.1002/mds.23429
- Troche, M. S., & Altmann, L. J. P. (2012). Sentence production in Parkinson disease: Effects of conceptual and task complexity. *Applied Psycholinguistics*, 33(02), 225-251.
- Tröster, A. I., Stalp, L. D., Paolo, A. M., Fields, J. A., & Koller, W. C. (1995). Neuropsychological impairment in Parkinson's disease with and without depression. *Archives of neurology*, 52(12), 1164-1169.
- Ulatowska, H. K., Allard, L., & Chapman, S. B. (1990). Narrative and procedural discourse in aphasia. In *Discourse ability and brain damage* (pp. 180-198). Springer.
- Ulatowska, H. K., Chapman, S. B., Bloom, R. L., & Obler, L. K. (1994). Discourse macrostructure in aphasia. *Discourse analysis and applications: Studies in adult clinical populations*, 29-46.
- Ulatowska, H. K., North, A. J., & Macaluso-Haynes, S. (1981). Production of narrative and procedural discourse in aphasia. *Brain and Language*, 13(2), 345-371.
- Ulatowska, H. K., & Olness, G. (2004). Discourse. In R. D. Kent (Ed.), *The MIT Encyclopedia of Communication Disorders* (pp. 300-302). Cambridge, MA: The MIT Press.
- Ullman, M. T. (2006). Is Broca's area part of a basal ganglia thalamocortical circuit? *Cortex*, 42(4), 480-485.
- United States Army Battery. (1944). *Manual of Directions and Scoring*. Washington,

- DC: War Department, Adjutant General's Office.
- van Laar, T. (2007). Parkinson's disease-related Pharmacotherapy. In E. C. Wolters, T. van Laar, & H. W. Berendse (Eds.), *Parkinsonism and Related Disorders* (pp. 219-239). Amsterdam: VU University Press.
- van Rooden, S. M., Colas, F., Martinez-Martin, P., Visser, M., Verbaan, D., Marinus, J., . . . van Hilten, J. J. (2011). Clinical subtypes of Parkinson's disease. *Movement Disorders*, 26(1), 51-58. doi:10.1002/mds.23346
- Verbaan, D., Marinus, J., Visser, M., van Rooden, S. M., Stiggelbout, A. M., Middelkoop, H. A., & van Hilten, J. J. (2007). Cognitive impairment in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, 78(11), 1182-1187. doi:10.1136/jnnp.2006.112367
- Webster, D. D. (1979). *Webster Rating Scale, Parkinson's Disease Patient Evaluation*. New York: Endo Laboratories.
- Whitworth, A., Lesser, R., & McKeith, I. A. N. (1999). Profiling conversation in Parkinson's disease with cognitive impairment. *Aphasiology*, 13(4-5), 407-425.
- Wolters, E. C., & Bosboom, J. L. W. (2007a). Parkinson's Disease-Related Cognitive Impairment and Dementia (PDD and DLB). In *Parkinsonism and Related Disorders* (pp. 191-201). Amsterdam: VU University Press.
- Wolters, E. C., & Bosboom, J. L. W. (2007b). Parkinson's Disease. In E. C. Wolters, T. van Laar, & H. W. Berendse (Eds.), *Parkinsonism and Related Disorders* (pp. 143-158). Amsterdam: VU University Press.
- Woods, S. P., Scott, J. C., Sires, D. A., Grant, I., Heaton, R. K., & Troster, A. I. (2005). Action (verb) fluency: test-retest reliability, normative standards, and construct validity. *Journal of the International Neuropsychological Society*, 11(4), 408-415.
- Wooten, G. F., Currie, L. J., Bovbjerg, V. E., Lee, J. K., & Patrie, J. (2004). Are men at greater risk for Parkinson's disease than women? *Journal of Neurology, Neurosurgery & Psychiatry*, 75(4), 637-639.
- Wright, H. H., Koutsoftas, A. D., Capilouto, G. J., & Fergadiotis, G. (2014). Global coherence in younger and older adults: Influence of cognitive processes and discourse type. *Aging, Neuropsychology, and Cognition: A Journal of Normal and Dysfunctional Development*, 21(2), 174-196. doi:10.1080/13825585.2013.794894
- Yorkston, K., Beukelman, D., & Traynor, C. (1984). *Assessment of Intelligibility of Dysarthric Speech*. Austin, TX: Pro-Ed.
- Yorkston, K., Beukelman, D., & Hakel, M. (1996). *Sentence Intelligibility Test for Windows*. Lincoln, NE: Institute for Rehabilitation Science and Engineering at Madonna Rehabilitation Hospital.
- Yorkston, K. M., & Beukelman, D. R. (1980). An analysis of connected speech samples of aphasic and normal speakers. *Journal of Speech and Hearing Disorders*, 45(1), 27-36.
- Zec, R. F., Burkett, N. R., Markwell, S. J., & Larsen, D. L. (2007). A cross-sectional study of the effects of age, education, and gender on the Boston Naming Test. *The Clinical Neuropsychologist*, 21(4), 587-616.
- Zgaljardic, D. J., Foldi, N. S., & Borod, J. C. (2004). Cognitive and behavioral dysfunction in Parkinson's disease: neurochemical and clinicopathological contributions. *Journal of Neural Transmission*, 111(10-11), 1287-1301.
- Zraick, R. I., Carr, P. B., Gregg, B. A., & Smith-Olinde, L. (2011). Information Units

Produced by Persons with Mild Alzheimer's Disease During a Picture Description Task. *Journal of Medical Speech-Language Pathology*, 19(2), 37-45.

Appendix A

Western Research Ethics Approval



**Western
Research**

Research Ethics

Use of Human Participants - Ethics Approval Notice

Principal Investigator: Dr. Joseph B. Orange
 File Number:103718
 Review Level:Delegated
 Approved Local Adult Participants:36
 Approved Local Minor Participants:0
 Protocol Title:Discourse Performance in Individuals with Parkinson Disease
 Department & Institution:Health Sciences|Communication Sciences & Disorders,Western University
 Sponsor:
 Ethics Approval Date:June 10, 2013 Expiry Date:December 31, 2014
 Documents Reviewed & Approved & Documents Received for Information:

Document Name	Comments	Version Date
Instruments	Appendix E: Standardized Language & Cognitive Measures	2013/04/06
Other	Appendix M: Failed hearing screen letter	2013/04/08
Letter of Information & Consent	Appendix L: Consent Form	2013/04/08
Instruments	Appendix B: Study Overview	2013/04/06
Instruments	Appendix F: Speech Intelligibility Protocol	2013/04/06
Other	Appendix N: Personal Amplifier Technical Specs	2013/04/06
Instruments	Appendix D: Hearing Screening	2013/04/06
Instruments	Appendix I: Forms for protocol	2013/04/06
Instruments	Appendix H: Motor/Disease Severity	2013/04/06
Instruments	Appendix G: Discourse Sampling	2013/04/06
Instruments	Appendix A: Power Calculation	2013/04/06
Instruments	Appendix C: Cognitive/Dementia Screening	2013/04/06
Western University Protocol		2013/04/09
Revised Caregiver Letter of Information & Consent	Revised LOI controls - version date added and issue of referral to attending MD re: cognitive screen score	2013/05/10
Revised Letter of Information & Consent	Revised LOI Participants with PD - version date added and issue of referral to attending MD re: cognitive screen score	2013/05/10
Revised Western University Protocol	Revised protocol - issue of referral to attending MD re: score on cognitive screen	2013/05/31

This is to notify you that The University of Western Ontario Research Ethics Board for Health Sciences Research Involving Human Subjects (HSREB) which is organized and operates according to the Tri-Council Policy Statement: Ethical Conduct of Research Involving Humans and the Health Canada/ICH Good Clinical Practice Practices, Consolidated Guidelines, and the applicable laws and regulations of Ontario has reviewed and granted approval to the above referenced revision(s) or amendment(s) on the approval date noted above. The membership of this REB also complies with the membership requirements for REB's as defined in Division 5 of the Food and Drug Regulations.

The ethics approval for this study shall remain valid until the expiry date noted above assuming timely and acceptable responses to the HSREB's periodic requests for surveillance and monitoring information. If you require an updated approval notice prior to that time you must request it using the University of Western Ontario Updated Approval Request Form.

Members of the HSREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the HSREB.

The Chair of the HSREB is Dr. Joseph Gilbert. The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Note. This form has been modified from its original format per university regulations

Appendix B

Lawson Health Research Institutes Approval



LAWSON FINAL APPROVAL NOTICE

LAWSON APPROVAL NUMBER: R-13-530

PROJECT TITLE: Discourse Performance in Individuals with Parkinson's Disease

PRINCIPAL INVESTIGATOR: Dr. Joseph Orange

LAWSON APPROVAL DATE: December 19, 2013

Health Sciences REB#: 103718

Please be advised that the above project was reviewed by the Clinical Research Impact Committee and Lawson Administration and the project:

Was Approved

Please inform the appropriate nursing units, laboratories, etc. before starting this protocol. The Lawson Approval Number must be used when communicating with these areas.

Dr. David Hill
V.P. Research
Lawson Health Research Institute

Note. This form has been modified from its original format per university regulations

Appendix C

Complete discourse protocol from which the tasks in the current study are abstracted. This document is reproduced from the ethics application made by Orange, Roberts and Jog approved June 10th, 2014.

Protocol:

- A total of eight discourse samples will be collected.
- To ensure a reliable sample for discourse analyses, two discourse samples will be collected for each genre (Doyle et al., 2000; Nicholas & Brookshire, 1993).
- The discourse genres to be sampled include: Story re-telling, story generation, procedural narrative, and personal narrative.
- The order of discourse tasks/genres will be randomized for each participant.
- Discourse samples will be audio recorded.
- Participants are allowed as much time as needed to produce discourse samples (no time limit to productions). Experience suggests that each discourse sample will average 3 minutes in length.
- To ensure validity and reliability of samples, standardized and published stimuli and protocols will be used to elicit each of the discourse genres (see below).
 - A) Story Re-Telling: Two stories from the Discourse Comprehension Test-2nd edition (Brookshire and Nicholas, 1997) will be used to elicit story re-telling discourse samples. These two stories are matched for length, complexity, and memory demands and have demonstrated good inter-story reliability (Doyle et al., 2000). Participants will listen to each story once at a comfortable volume. After participants finish listening to the story they will be prompted by the researcher to re-tell the story in as much detail as they can remember.
 - B) Story Generation: Two standardized picture sequences (6 pictures/sequence) from Nicholas and Brookshire, 1993 will be used to elicit story generation discourse samples. These two stimuli are well-established in the field of discourse studies. Participants will be asked to review the pictures and then will be prompted by the researcher to tell a story about the events in the picture sequence.
 - C) Procedural: Two standardized elicitation stimuli will be used to generate procedural discourse samples (Nicholas and Brookshire, 1993). Participants will be asked to describe in detail the procedure for completing two tasks. These two

stimuli are well-established in the field of discourse studies. One stimuli request is: "Tell me how you would go about washing dishes by hand" (Nicholas and Brookshire, 1993). The other is "Tell me how you would go about writing and sending a letter" (Nicholas and Brookshire, 1993).

D) Personal Narrative: Two standardized elicitation stimuli will be used to generate personal narrative discourse samples (Nicholas and Brookshire, 1993). These two stimuli are well-established in the field of discourse studies. The two stimuli requests are "Tell me what you usually do on Sundays" and "Tell me where you live and describe it to me" (Nicholas and Brookshire, 1993).

Appendix D

CHAT Symbols Used in Coding Discourse Transcripts

Symbol	Coding Reference	Examples
.	Period. End of C-Unit	<i>He reached for the cookies.</i>
?	Question. End of C-Unit	<i>Did he reach for the cookies?</i> Upward inflection on <i>cookies</i> denotes question
!	Exclamation. End of C-Unit	<i>Oh boy!</i> Inflection denotes excitement
+...	Trailing off/incomplete utterance	*AR: <i>He went +...</i> *AR: <i>Did you know today is Tuesday?</i>
+”/.	Quotation follows next line when quotation = separate C-Unit	*AR: <i>The wife said to the husband [“] I have had enough.” +”/.</i> *AR: <i>[“] You should really do more around the house.”</i>
[“]	Denotes character dialogue or direct quote follows	*AR: <i>She said [“] I’m leaving.”</i>
„	Denotes a tag question	*AR: <i>She left for her mother’s „, didn’t she?</i>
#	1 to 2 second pause between words	<i>Well uh # the wife is angry at the man.</i>
[# X.0]	Pauses longer than 2 seconds with X denoting the actual duration	<i>Well uh [# 4.0] the wife is angry at the man.</i>
:	Lengthened syllable	<i>Well I see a man and a woman argu:ing.</i>
~	Pauses between syllables within a word	<i>She ran the car into the tree out~side.</i>
[/]	Retracing without correction. Repetition.	< <i>She packed</i> > [/] <i>she packed her suitcase.</i> Phrase retracing/repetition
[//]	Retracing with correction. Revision.	< <i>She packed</i> > [//] <i>she picked up her suitcase</i>
XXX	Unintelligible spoken language (more than one word)	<i>She packed her bags and XXX in a hurry.</i>
XX	Unintelligible spoken language (single word)	<i>She XX her bags and ran out the door in a hurry.</i>

[?]	Best guess at what was said	<i>She ran [?] out the door in a hurry.</i>
()	Non completion of a word	<i>She ran out the d() in a hurry.</i>
0word	Omitted word	<i>She 0word out the door in a hurry.</i>

Note. Coding symbols were adapted and reprinted with permission of the author and CHILDES) from *The CHILDES Project Tools for Analyzing Talk – Electronic Edition Part 1: The CHAT Transcription Format.* (MacWhinney, 2014)

Appendix E

Conventions for Segmenting C-units (SALT, LLC). Used and reprinted with permission of SALT, LLC.

Summary of C-unit Segmentation Rules

Definitions

- C-Unit - independent clause with its modifiers. It includes one main clause with all subordinate clauses attached to it. It cannot be further divided without the disappearance of its essential meaning.
- Clause - statement containing both a subject and a predicate.

Segmenting Utterances into C-Units

Main clauses can stand by themselves and can be segmented into one C-unit. Subordinate clauses **DEPEND** on the main clause to make sense. They cannot stand alone or be separated from the main clause. So a C-unit will either consist of a main clause or a main clause with its subordinating clause(s).

Coordinating and Subordination Conjunctions

• Coordination Conjunctions

The segmenting rule is simple when utterances contain coordinating conjunctions. These conjunctions link two **main** clauses which should be separated/segmented into two utterances (or two C-units) that can each stand alone. Common coordinating conjunctions include: and, but, so (but not "so that"), and then, then.

Example 1:

C The frog was sit/ing on a lily pad.

C And then it jump/ed in.

Example 2:

C He had to catch the frog.

C Or the waiter would make them leave.

• Subordinating Conjunctions

Subordinating conjunctions **link** a main clause and a subordinate clause. A C-unit includes the main clause with all subordinate clauses attached to it. The following are examples of subordinating conjunctions:

Early Development: because, that, when, who

Later Development: after, before, so (that), which, although, if, unless, while, as, how, until, as __as, like, where, since, although, who, before, how, while

Examples:

C When the boy saw it, the frog jump/ed.

C We can/'t find my cat who always run/3s away.

• "because" and "so"

Always consider "because" as a subordinating conjunction (it will never begin an utterance) unless it is preceded by the utterance of another speaker.

The word "so" can be either a coordinating conjunction or a subordinating conjunction. If its usage means "so that", it is a subordinating conjunction. Otherwise it is a coordinating conjunction.

Other rules for segmenting C-units

• Sentence fragments

Sentence fragments are counted as separate C-units when the final intonation contour of the utterance indicates that a complete thought has been spoken. For example:

C The boy, the dog, and the frog, they were friend/s.

versus

C The boy, the dog, and the frog. { *fragment based on intonation* }

C They were friend/s.

• Elliptical responses

Elliptical responses (*sentence fragments*) to questions or prompts from the examiner are counted as separate C-units. For example:

E What did you do next?

C Shop/ed.

- **Yes/No responses**

Separate the yes/no response from the utterance which follows it. For example,

- E Do you want to read your book now?
- C No.
- C I don't.

- **Tags**

Do not segment phrases such as "you know", "I guess", and "I mean" when they are used as tags. For example,

- C He's gonna live with his dad, I guess.
- C And then, you know, they were go/ing to this town.

- **Questions as tags**

Do not segment questions when they are used as tags. For example,

- C They got in trouble, right?
- C He miss/ed the bus, did/n't he?
- C That movie was good, would/n't you agree?

- **Dialogue Complement/Complement**

Dialogue quotes which are embedded in, or as part of, an utterance are counted as one C-unit as in this example:

- C And he call/ed, "Where are you frog"?
- C And then the boy said, "Be quiet".

Successive main clauses that occur in dialogue quotes are counted as separate C-units. For example,

- C And he said, "I'm ready".
- C "I want to go to the store now".

Complement:

- C She thought, Sam was incorrect.
- C He realized, nothing has changed

- **Grammatical errors**

Ignore grammatical errors when segmenting utterances. For example,

- C They is[EW:are] go/ing now. { child said, "They is going now." }
- C We *are go/ing too. { child said, "We going too." }

- **Pauses and intonation**

Do not ignore pauses and intonation when segmenting utterances but, whenever reasonable, segment utterances based on grammar rules. When listening to speech, for example, there is sometimes a significant pause (with or without ending intonation) between a main clause and a subordinate clause. This inclines one to segment the utterance. With C-unit segmentation, however, the utterance would not be segmented as in this following example where the speaker paused for two seconds between the main clause and the subordinate clause:

- C I like/ed the movie alot :02 because it was really funny.

In the following segment, however, you have to consider pause time and intonation:

- C I like/ed the movie alot.
- : :02
- E Mhm.
- C Because it was really funny.

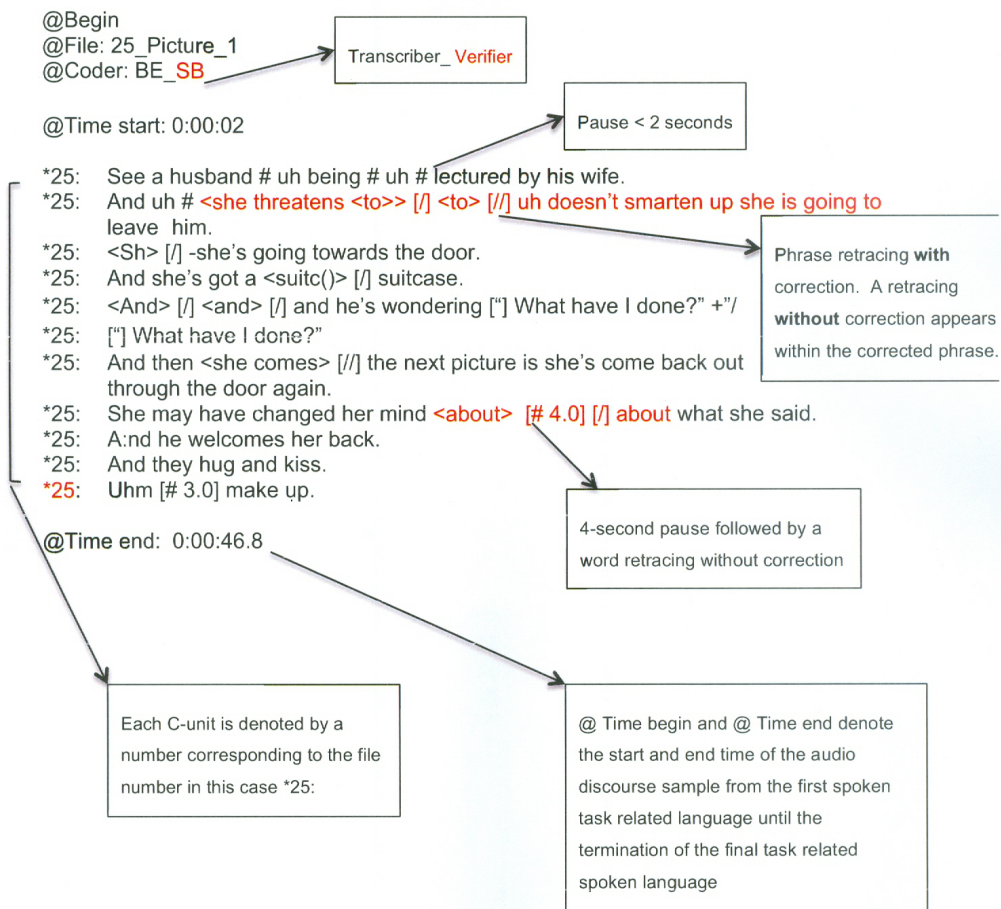
If there is a significant pause and ending intonation (falling for statements, rising for questions) between the speaker's first utterance and the examiner's "Mhm", segment the utterances as show above. Otherwise, give the speaker credit for subordination and transcribe these "prompt sounds" as interjections as follows:

C I like/ed the movie alot :02 < > because it was really funny.
E <Mhm>.

Note. Materials reproduced with permission of the author SALT Software, LLC retrieved from <http://www.saltsoftware.com/resources/tran aids/index.cfm#>.

Appendix F

Sample of an Orthographically Transcribed, Coded, and Segmented Audio File Prepared for Discourse Analyses



Appendix G

Discourse Analysis Data Form
Angela Roberts Thesis (2014)

Sample # Click here to enter text.
text.

Researcher Click here to enter text.

Item	Analysis Measure	Value
1	Total words	
2	Total Duration with Pauses	
3	Total C-Units	
4	Number of words/C-unit	
5	Words/Minute ((#1 ÷ #2)*60)	
6	# of Grammatically Correct C-Units	
7	% Grammatically Correct C-Units (#6 ÷ #3)	
8	# of C-Units with complex grammatical structures	
9	% of C-Units with complex grammatical structures (#8 ÷ #3)	
10	# CIUs	
11	% CIUs (#10 ÷ #1)	
12	CIUs/Minute (#10 ÷ #2)	
13	Total Open Class Words	
14	% Open Class Words (#13 ÷ #1)	
16	Total Verbs	
17	% of Verbs (#17 ÷ #1)	
<i>Verbal Fragmentation Analysis</i>		
18	Total Verbal Disruptions	
19	Verbal Disruptions/100 Words (#20 + # 23 + #27 + #28 + #29 + #30) per 100 words	
20	Total Unfilled pauses > 2 seconds	
21	Word revision	
22	Phrase revision	
23	Total Revisions (#21 + #22)	
24	Phrase repetition	
25	Word repetition	
26	Sound/Syllable repetition	
27	Total Repetition (#24 + #25 + #26)	
28	Total Filled Pauses (non-word)	
29	Total Incomplete utterance	
30	Total Empty speech (Nicholas et al., 1985)	

Appendix H

Discourse Analyses Definitions/Procedures (Rzepczyk, 2001)

Total Words

Words are defined and quantified using Nicholas and Brookshire's (1993) definitions and guidelines. According to their definition, words are "intelligible in context to someone who knows the topic being discussed". They "do not have to be accurate, relevant, or informative relative to the topic being discussed" (p.348). The following rules applied:

- Interjections (e.g., oh, wow, golly, gosh, gee, aha, ahh, shhh) are counted as one word
- Informal terms (e.g., uh-huh [affirmative], un-uh [negative], nope, yep, yeah, mhmm) are counted as one word.
- Common contractions and simplifications (e.g., "gonna" for "going to", "em" for "them", 'n' or an' for "and") are counted. The quantity assigned depends on the original form (i.e. "gonna" for "going to" would be 2 words; "em" for "them" would be 1 word).
- Standard contractions (e.g., don't, he's) and colloquial contractions (e.g., gonna, sorta) are counted as two words.
- Each word in hyphenated forms is counted as one word (e.g., jack-in-the-box = 4 words).
- Each word in numbers is counted as one word (e.g., one hundred thirty-four = 4 words).
- Compound words (e.g., pancake, cowboy) are counted as one word.
- If a pause occurs in the middle of a compound word, it is still counted as only one word (e.g., shuffle [pause] board = 1 word).
- If a filler occurs in the middle of the compound word, then it is counted as two words (e.g., sun uh room = 2 words).
- If a revision or repetition occurs within a compound word, it is still counted as one word (e.g., sunr-room = 1 word).
- Each word in proper names is counted as one word (e.g., Mary Smith = 2 words).
- Acronyms (e.g., TWA, GM) are counted as one word.
- Initials (e.g., K.G.) are counted as one word.
- Words in revisions and repetitions are counted as separate words (i.e., both the words in original and the words in the revised portions of the phrase/utterance are counted. "he went //no she went//" = 5 words)
- Single letters are not counted (e.g., N6H = 1 word ['six']).
- Non-word fillers (e.g., um, er, uh, hmm, mmm) are not counted.
- Paraphasic errors that result in non-words (e.g., crolesterol) are not counted.

Duration

The duration of the discourse sample was calculated by subtracting the ending time from the starting time as denoted on the discourse transcript. The @ Time start: was defined as the initiation of the first word of the task related spoken language and the @ End time: was defined as the termination of the final word of the task related spoken language in the discourse audio file.

Total C-Units

The transcripts were segmented using the SALT conventions for C-units during the transcription process. Each new C-unit is denoted by *## with '##' = the sample number (e.g., 01 = sample 1). The total number of C-units equals the number of separate C-units as noted above.

Words per Minute (WPM)

Divide the total number of words by duration (measured in seconds) and then multiply by 60 to generate WPM.

Number of Words per C-unit (Words/C-unit)

Divide the total number of words by the total number of C-units.

of Grammatically Well-Formed C-Units

Count the total number of C-Units that meet the following criteria:

A grammatically incorrect utterance is defined as an utterance that violates one or more grammar rules, as defined by Quirk and Greenbaum (1973).

The following conventions apply:

- Subject-verb concord of number and person (e.g., The window is open.).
- Subject-complement concord (e.g., The child was an angel.).
- Subject-object concord. This structure applies where the second element is a reflexive pronoun (e.g., He injured himself in the leg.).
- Pronoun concord (e.g., Jeff finished his homework.; Jeff and Lisa finished their homework.).
- Temporal concord (e.g., Yesterday I went to the store.).
- Correct word order.
- Labeling in the picture description task is acceptable as grammatical unless an
 - obligatory article is omitted (e.g., // car // = grammatical; // and car // = grammatically incorrect).
- Ellipsis is scored as grammatical.
- The use of the verb “got” in place of the verb “have” (e.g., I got three children.) is scored as grammatical.
- Only complete utterances are scored according to grammaticality. Utterances that are comprised of only an affirmative (e.g., yeah mmmm, sure), or negative (e.g., no) response, or a one-word request for clarification (e.g., huh, what) are not scored for grammaticality.

- Syntactically correct but semantically anomalous utterances are scored as grammatical.
- Repetitions and all but the final revision of an utterance are excluded from this analysis.

% Grammatically Well-Formed C-units (% Grammatical)

Divide the total number of grammatically well-formed C-units by the Total number of C-units.

of C-Units with Complex Grammar Structures

C-Units with complex grammatical structures are defined by the criteria used by Ash et al. (pp. 381-382) (Ash et al., 2009). C-units that contain either or both of dependent clauses or phrasal adjuncts are defined as C-Units with complex grammatical structures:

- **Dependent clauses** defined as a phrase containing a subject and a verb that cannot stand alone as an independent sentence. Dependent phrases are typically introduced by subordinating conjunctions (e.g., while, when, because, than) or by a complementizer (e.g., that, as in “The woman thought that the children were in their room”). The phrase must contain a subject and an inflected verb to constitute a clause.
- **Phrasal adjuncts** defined as a phrase that is outside the subject noun phrase and outside the verb phrase but the content applies to the sentence as a whole. Phrasal adjuncts are commonly introduced by an infinitive verb or gerund. (e.g., “The went outside to see if they could find a four-leafed clover” – ‘to see if they could find.... = a phrasal adjunct.

% of C-Units with Complex Grammatical Structures (% Complex

Divide the total number of C-Units containing complex grammatical structures by the Total number of C-Units.

Total Open Class Words

Open Class Words. Open class words are defined according to the definition of Saffran et al. (1989) with some modifications. The following conventions apply:

- Nouns are counted.
- Main verbs are counted.
- “Be, do, have count as open class when they occur as a main verb. Going to, have to, etc. count as closed class when used as auxiliary verbs as substitutes for will and must, respectively” (p.472).
- Adverbs are counted.
- Adjectives are counted.
- “Numerals are considered open class words. The only exception is ‘one’ when used as a pronoun (e.g., “I saw one on the table”).
- Note that one, when modified by a determiner, is counted as open class: “Cinderella was the one.” (p. 472)
- Open class words occurring in revisions are counted.
- Only the first occurrence of open class words in repetitions are counted.

Percentage of Open Class Words (% Open)

The total number of open class words divided by the Total number of words

Total Verbs

Verb. The following conventions apply:

- Participles (e.g., present: verb + ing; past: verb + ed) are counted, except when they are used as adjectives (e.g., The movie was entertaining - “entertaining” = adjective; He was entertaining the audience - “entertaining” = verb).
- Infinitives (i.e., to + verb) are counted.
- Gerunds (i.e., verb form ending in -ing used as nouns, such as “I quit smoking.”) are counted.
- Verbs acting as auxiliaries (e.g., be, have, do) are not counted.
- Going to, have to, got to are not counted when used as auxiliary verbs as substitutes for will and must.
- Modals (e.g., can, could, may, would, shall, should, will, must, used to) are not counted.
- Verbs occurring in revisions are counted.
- Only the first occurrence of verbs in repetitions is counted.

Percentage of Verbs (% Verbs)

Divide the total number of verbs by the total number of words.

Total Verbal Disruptions

Total all occurrences of verbal disruptions in the transcript.

Number of Verbal Disruptions Per 100 Words (Disrupt/100 Words)

Total of all verbal disruptions divided by total words and multiplied by 100.

Subtypes of Verbal Disruptions Counted

Unfilled Pauses

Unfilled pauses refer to all within and between sentence pauses greater than or equal to 2 seconds duration (Illes, Metter, Hanson, & Iritani, 1988). This will be denoted in the transcript by [# X.0] where the exact value entered should be 2 seconds or greater.

Revisions

Self-correction/Revision (i.e., referred to by MacWhinney as ‘retracing’’) refers to all instances of “retracing with correction” as defined in the CHILDES Project (MacWhinney, 2000). The symbol for this in the transcript is [//]. According to this definition, self-corrections/revisions refer to instances when “a speaker starts to say something, stops, repeats the basic phrase, changes the syntax but maintains the same idea” (p.74). The following conventions apply:

- Revisions at the word, and phrase levels are counted separately and entered separately into the analysis form.
- Each revision is counted (e.g., I [we] [he] went = 2 self-revisions - word).
- A word produced alone and then followed by a contracted form is counted as a revision (e.g., she [she's] going = 1 revision - word).
- Part of a word or a phrase that is followed by its complete form is counted as a revision (e.g., tha-[that's] = 1 revision - word; she we- [she went] = 1 revision - word).
- The rule also applies to revisions within compound words (e.g., sunr-[room] = 1 revision).
- Revisions separated by non-word fillers are counted (e.g., I went uh [she went] to the store = 1 revision).
- Revisions separated by word fillers are counted (e.g., I went to the you know [that] place = 1 revision).
- Revisions that occur within revisions are counted (e.g., I took the bus [this [that] train] = 2 revisions).
- Revisions that occur within repetitions are counted (e.g., I went (I we- [went]) = 1 revision [] within 1 phrase repetition).
- Revisions of non-word fillers are **not** counted (e.g., uh um uh = 0 self-corrections).

Repetitions

Repetition refers to all instances of “retracing without correction” as defined in the CHILDES Project (MacWhinney, 2000). According to this definition, repetitions refer to instances when “a speaker begins to say something, stops, and then repeats the earlier material without change” (p.73). In the transcripts, episodes of repetition are denoted by [/]. The following conventions apply:

- Word, phrase and sound/syllable repetitions are counted separately.
- Each repetition is counted (e.g., He (he) (he) (he) fell = 3 word repetitions).
- Repetition of non-word fillers are not counted (e.g., uh uh uh = 0 repetitions).
- Repetitions separated by non-word fillers are counted (e.g., He um (he) fell. = 1 word repetition).
- A word produced alone and then followed by a contracted form is counted as a revision (e.g., she [she's] going) and not as a repetition.
- A phrase or word that is partially repeated following its complete form is counted as a repetition (e.g., that (tha-) = 1 part word repetition; she went (she we-) = 1 phrase repetition).
- Repetitions that occur within repetitions are counted (e.g., I know (I (I) know) = 1 word repetition and 1 phrase repetition).
- Repetitions that occur within revisions are counted (e.g., I went [she (she) went] = 1 word repetition and 1 revision).

Filled Pauses

Non-Word fillers (e.g., 'uh', 'uhm') used as a filler or a starter.

Incomplete Utterance/Trailing off

Incomplete utterances are defined following the definitions and guidelines of the CHILDE Project (MacWhinney, 2000). According to this definition, utterances are "incomplete but not interrupted" utterances in which the participant "shifts attention away from what they are saying, sometimes even forgetting what they were going to say" and is usually "followed by a pause" (p.64). The presence of an incomplete utterance is denoted in the transcript by the symbol +...

Empty Speech (adapted from Nicholas et al., 1985)

Empty speech is defined as words/phrases that do not contribute substantially to the content information of the sample. The following conventions apply:

- Empty phrases = words or phrases used as continuation devices (e.g., and so on, like that)
- Indefinite terms = nonspecific nouns (e.g., junk, stuff, thing)
- Deictic terms (e.g., here, there, this, that)
- Pronouns without antecedents
- Comments on task (e.g., that's hard, I don't know how)
- Literal/Phonemic paraphasia: Non-words or real words that are phonologically related to the target word (e.g., *tofa* for *sofa*; *red* for *bed*).
- Semantic paraphasia: Real words that are semantically related to the target word (e.g., *chair* for *stool*).
- Unrelated Verbal paraphasia: Real words with no semantic relation to the target word (e.g., *car* for *jacket*).
- Neologism: Non-words with no apparent relation to the target word (e.g., *chite* for *mother*).

Total Correct Information Units (CIUs or Total CIUs)

Use the attached definition and copyrighted materials from Nicholas and Brookshire (1993) (Appendix I). Materials used and reproduced with the explicit permission of Linda Nicholas the surviving author and copyright holder.

Percentage of CIUs (% CIUs)

Divide the total CIUs by the Total number of words.

CIUs per Minute (CIUs/Minute)

Divide the total CIUs by the duration of the discourse sample.

Appendix I

Rules for Calculating CIUs.

348 *Journal of Speech and Hearing Research*

36 338-350 April 1993

Appendix B

Rules for scoring and counting words and correct information units (CIUs)

Prior to determining which words should be included in counts of words and correct information units, delete statements that are made before or after the speaker performs the task or suggest that the speaker is ready to begin or has finished the task and do not provide information about the picture(s) or topic itself. Such statements generally are not produced consistently by speakers from one session to another and are deleted to help stabilized counts across sessions.

- I hope I can remember how I did this before.
- I'll start by saying this.
- I'm supposed to tell you about washing dishes.
- I'm ready to start.
- That's about it.
- I can't say any more.
- The end.
- That's about what our Sundays are like.

These statements should be grammatically separate from discussion of the picture(s) or topic. The following first statements by a speaker would be included in the word count.

- In the first picture, the man is angry.
- Well first of all, there's a couple fighting.
- Okay, there's a man and a woman.
- Well now, here's a picture of a party.

This does not include commentary on the task or on the speaker's performance that occurs while the speaker is discussing the picture(s) or topic. (See 1.22 for rules about commentary.)

Instructions: Draw a horizontal line through the middle of words that are to be deleted prior to making decisions about the word count. (~~That's about it~~)

1.0. COUNTING WORDS

Definition: To be included in the word count, words must be intelligible in context to someone who knows the picture(s) or topic being discussed. Context refers to what the scorer knows about the picture(s) or topic and what the scorer knows from the speaker's prior words. Words do not have to be accurate, relevant, or informative relative to the picture(s) or topic being discussed to be included in the word count.

Instructions: Cross out with red Xs words that are not to be included in the word count.

RULES FOR COUNTING WORDS

1.1. DO NOT COUNT THE FOLLOWING

1.11. Words or partial words that are not intelligible in context to someone who knows the picture(s) or topic being discussed.

- He went to the ~~framp~~.
- That appears to be a ~~norble~~.
- He had a ~~st~~ . . . an . . . steak.

1.12. Nonword filler (~~um~~, ~~er~~, ~~uh~~). (See 1.23 and 1.24 for a rule dealing with filler words and phrases, interjections, and informal terms.)

1.2. COUNT THE FOLLOWING

1.21. All words that are intelligible in context. Count words that contain sound substitutions, omissions, distortions, or additions if the word is intelligible in context (~~hiccup~~ for hiccup). If the incorrect production results in another real word that does not appear to be the target word, it is still included in the word count (~~paper~~ for pepper).

- 1.22. Commentary on the task, on the speaker's performance, or on the speaker's experiences.
- This is pretty hard.
- I can't think of that word.
- No, that's not right.
- My wife and I used to fight like that.

1.23. Filler words and phrases (~~you know~~, ~~I mean~~, ~~okay~~). Do not count nonword filler. (See 1.12.)

1.24. Interjections (~~oh~~, ~~oh boy~~, ~~wow~~, ~~golly~~, ~~gosh~~, ~~gee~~, ~~aha~~, ~~hmm~~) and informal terms (~~uh-huh~~ [~~affirmative~~], ~~un-uh~~ [~~negative~~], ~~nope~~, ~~yep~~, ~~yeah~~).

1.25. Common contractions or simplifications of words (~~gonna~~ for ~~going to~~, ~~sorta~~ for ~~sort of~~, ~~em~~ for ~~them~~). Contractions (both standard [~~don't~~, ~~he's~~] and colloquial [~~gonna~~, ~~sorta~~]) are counted as two words.

1.26. Each word in hyphenated words (~~jack-in-the-box~~ = 4 words).

1.27. Each word in numbers (~~twenty-two~~ = 2 words, ~~one hundred thirty-four~~ = 4 words, ~~nineteen fifty-five~~ = 3 words).

1.28. Compound words as one word (~~pancake~~, ~~cowboy~~).

1.29. Each word in proper names (~~Mary Smith~~, ~~St. Paul~~, ~~Mason City~~ = 2 words each).

1.30. Count acronyms as one word (~~VA~~, ~~VFW~~, ~~TWA~~ = 1 word each).

2.0. COUNTING CORRECT INFORMATION UNITS (CIUs)

Definition: Correct information units are words that are **intelligible** in context, **accurate** in relation to the picture(s) or topic, and **relevant** to and **informative** about the content of the picture(s) or the topic. Words do not have to be used in a grammatically correct manner to be included in the correct information count. Each correct information unit consists of a single word and only words that have been included in the word count can be considered for inclusion in the correct information unit count.

Instructions: Put a diagonal pencilled slash through words that are **not** to be included in the correct information count (~~man~~).

RULES FOR COUNTING CIUs

2.1. DO NOT COUNT THE FOLLOWING

(In this section, words in **bold print** would **not** be counted as correct information units.)

2.11. Words that do not accurately portray what is in the picture(s) or that do not seem accurate in relation to the topic being discussed, such as incorrect names, pronouns, numbers, actions, etc. If a word reflects regional usage (such as calling the midday meal "dinner" in some areas), it is counted as a correct information unit. If grammatical incorrectness would lead to misunderstanding or uncertainty about the meaning of words, the grammatically incorrect words would not be counted as correct information units. (See 3.12 for examples of grammatically incorrect words that would be counted as correct information units.)

- The girl ~~is~~ riding her bike. (The picture shows a girl with a bike nearby which she may have been riding, but which she is not currently riding.)
- The girl is on a ~~ladder~~. ~~She fell~~. (The picture shows a boy on a stool who is tipping but has not fallen yet.)
- The ~~boys~~ and ~~girls~~ are arriving. (The picture shows only one boy and one girl arriving.)

If several people are involved in an action and only one of them is mentioned, the mentioned one is still counted as a correct information unit. This constitutes an incomplete description but not an inaccurate one.

The boy is arriving. (The picture shows a boy and a girl arriving.)

The man drove away. (The picture shows a couple driving away.)

2.12. Attempts to correct sound errors in words except for the final attempt.

- He put ~~paper~~ ~~popper~~ pepper on his food.
- She saw her with her ~~mass~~ . . . ~~mack~~ . . . mask.
- 2.13. Dead ends, false starts, or revisions in which the speaker begins an utterance but either revises it or leaves it uncompleted and uninformative with regard to the picture(s) or topic.
- My ~~si~~ . . . ~~no no not~~ my sister . . . my ~~fa~~ . . . with my wife.
- He goes over to her and ~~puts his~~ wants to give her a hug.
- He looks out and sees that ~~she had~~ the car ran into the tree.
- The . . . the . . . that one ~~oh~~ forget it.

- In the house in the mouse in the house
 - We go to a party no I mean a movie
- If an utterance is incomplete, but some information about the picture(s) or topic has been given, count that information.

• The kitchen window was . . .

In this example, the words the kitchen window was would be counted as correct information units (if they meet the other criteria). Even though the entire statement was not completed, the words are informative.

Words that express some legitimate uncertainty or change in perception about characters, events, or settings in a picture are counted as correct information units (if they meet the other criteria). See 2.18 for further examples.

- Her dad or maybe a neighbor was in the tree.
- From the looks of the candles, he must be four. No there is another candle on the table so he must be five years old.

2.14. Repetition of words or ideas that do not add new information to the utterance, are not necessary for cohesion or grammatical correctness, and are not purposely used to intensify meaning.

- The blue truck was blue.
- The restaurant was a new one. It was a new restaurant.
- She was cleaning washing the dishes.

Such repetition of words or ideas can be separated by other counted words.

- The mother was very angry. The daughter was crying. The mother was very mad.

Exceptions:

(a) If the repeated words or ideas are necessary for cohesion, they are counted. • She went to the store. The store was closed.

(b) If words are repeated to achieve effect or to intensify a statement they are counted.

- The girl was very, very sad.
- They were fighting, really fighting.

(c) If repeated words are used to expand on previous information, they are counted.

- He put on a shoe . . . a left shoe.
- There were some people . . . a man and a woman.

2.15. The first use of a pronoun for which an unambiguous referent has not been provided. Subsequent uses of the pronoun for the same unspecified or ambiguous referent are counted as correct information units (if they meet the other criteria).

- She (no referent) was doing the dishes. I think she was daydreaming.

If an inaccurate referent is provided but it is clear that a pronoun refers back to it, the pronoun would be counted as a correct information unit.

- The fox (inaccurate referent) ate some of the cake and it was hiding.

2.16. Vague or nonspecific words or phrases that are not necessary for the grammatical completeness of a statement and for which the subject has not provided a clear referent and for which the subject could have provided a more specific word or phrase.

- The mother is drying one of those things.
- She gave him some stuff.
- He put something up to the tree but that one knocked it down.
- We had pancakes or scrambled eggs or something like that.
- I wash the glasses and plates and so on.

The words "here" and "there" frequently fall into this category.

- Here we have a boy.
- This here boy is crying.
- That mother there is doing dishes.
- There is a cat here and a dog there.
- The mother is there.
- She put them over here.
- She has a bike there.
- The cookies were up there.

The following are examples of uses of "here" and "there" that are necessary for the grammatical completeness of the statement and cannot be replaced by a more specific word. These uses of "here" and "there" would be counted as correct information units.

- There is a boy.
- Here comes the same couple.

The following is an example of a nonspecific word that is preceded by a clear referent and would be counted as a correct information unit.

- The boy opened the cupboard. The cookies were up there.
- 2.17. Conjunctive terms (particularly **so** and **then**) if they are used indiscriminately as filler or continuants rather than as cohesive ties to connect ideas.

• There is a man. **Then** there is a woman and **then** a cat.

When used cohesively, "then" indicates the temporal order or sequential organization of things or events.

- She had lunch and then she went to the store.
- When you go into my house you see the living room first, then the dining room, then the kitchen.

When used cohesively, "so" indicates a casual consequence.

- He was thirsty so he drank some juice.
- The mother was after the dog so the boy was crying.

2.18. Qualifiers and modifiers if they are used indiscriminately as filler or are used unnecessarily in descriptions of events, settings, or characters that are unambiguously pictured. The following examples concern unambiguously pictured information.

- Apparently this is a kitchen.
- Evidently the boy is on a stool.
- I think that the cat is in the tree.
- It looks like the man is up in the tree too.
- The boy is sort of crying and the dog is kind of hiding.
- Of course, the woman left in a huff.

When used informatively, qualifiers and modifiers suggest legitimate uncertainty on the part of the speaker about events, settings, or characters portrayed in the picture(s) or modify associated words in a meaningful way. The following examples concern ambiguously pictured information.

- Apparently this is a mother and her two children.
- I think she is his sister.
- It looks like he gave them the wrong directions.
- She must be daydreaming.
- He might be the girl's dad or maybe he's a neighbor.
- He is the father or a neighbor. I don't know which.
- He looks sort of sad.
- Evidently they went around in a circle.

2.19. Filler words and phrases (you know, like, well, I mean, okay, oh well, anyway, yeah), interjections when they do not convey information about the content of the picture(s) or topic (oh, oh boy, wow, gosh, gee, golly, aha, hmmm), and tag questions (It is really smashed up, isn't it).

2.20. The conjunction "and." "And" is never counted as a correct information unit because it is often used as filler and we have found that its use as filler cannot be discriminated reliably from its uses as a conjunction.

2.21. Commentary on the task and lead-in phrases that do not give information about the picture(s) or topic and are not necessary for the grammatical completeness of the statement.

- These pictures are poorly drawn.
- This is kind of hard.
- In the first picture . . .
- As I said the last time, she was upset.

2.22. Commentary on the subject's performance or personal experiences.

- I can't think of the name of that.
- I can't say it.
- No, that's not right.
- My kids were always getting into trouble too.
- My wife and I used to fight like that.
- They are fighting but I don't know why.

Some statements that contain personal information may be appropriate in procedural and personal information descriptions and, in such cases, they would be counted as correct information units (if they meet the other criteria).

See 3.16 for embellishments that are counted as correct information units.

See previous page for statements that are deleted before beginning the word and correct information unit counts.

3.1. COUNT THE FOLLOWING (if they meet all other criteria)
(In this section, words in bold print would be counted as correct information units.)

3.11. All words (nouns, adjectives, pronouns, verbs, adverbs, articles, prepositions, and conjunctions) that are intelligible in context, accurate in relation to the picture(s) or topic, and relevant to and informative about the content of the picture(s) or topic.

3.12. Words do not have to be used in a grammatically correct manner to be counted. Words that violate standard English grammar rules concerning appropriate verb tense and form, agreement in number between subject and predicate, agreement between articles and nouns, incorrect use of articles, and appropriate singular and plural forms are counted as correct information units unless these violations would lead to misunderstanding or uncertainty about the meaning of the words.

See 2.11 for examples of words that would not be counted as correct information units.

- The **firemans** are coming.
- The firemen **ain't** rescued them yet.
- Put some stamp on it.
- The friends **is** here.
- He **don't** look very happy.

3.13. Production of a word that results in another English word, if the production would be intelligible as the target word in context.

- He is standing on a **school** and it is tipping over.

3.14. The final attempt in a series of attempts to correct sound errors.

- He went to the musket . . . minuet . . . **market**.

3.15. Informal terms (**nope**, **yep**, **uh-huh**, **un-uh**) when they convey information about the content of the picture(s) or topic.

- She said "Uh-huh, I'll do it."

3.16. Words in embellishments that add to the events portrayed in the picture(s) or express a moral, if they are consistent with the situation or events portrayed. Words that express some legitimate uncertainty about characters, settings, or events in the pictures.

- **He's going to get hurt and his mom is going to be angry.**
- **Some days everything seems to go wrong.**
- **That looks like a nice way to spend a summer day.**
- **Sooner or later cats usually get stuck up a tree.**
- **Mothers sometimes get distracted and don't notice things.**
- **This is the one about the accident-prone family.**

However, see 2.22 for examples of extraneous commentary that may resemble embellishments, but are not counted.

3.17. Verbs and auxiliary verbs (**is**, **are**, **was**, **were**, **to**, **has**, **have**, **will**, **would**, **has been**, etc.) as two separate correct information units—one for the auxiliary verb and one for the main verb.

- His mom **is going to be angry**. (Each word in bold print is a correct information unit.)

3.18. Contractions [both standard (**won't**) and colloquial (**gonna**)] as two correct information units.

3.19. Each word in hyphenated words (**father-in-law**, **good-bye**).

Copyright 1992 by Linda E. Nicholas and Robert H. Brookshire

Note. Materials reproduced from Nicholas and Brookshire (1993) with permission of the authors and publisher.

VITA

Angela Christine Roberts

Post-Secondary Education and Degrees

University of Western Ontario
Health and Rehabilitation Sciences
Speech and Language Sciences field
2009 – current; PhD(c)

University of Tennessee Knoxville
Speech Pathology
1995; MA

University of Tennessee Knoxville
Speech Pathology
1993; BA

Awards and Honors

Parkwood Hospital Endowed Fellowship in the Care of the
Elderly
2014-2015

Queen's Jubilee Medal of Honour for service to Canadians
living with Parkinson disease
2013

Western University Teaching Award
2013

Canadian Institutes of Health Research Fellowship Award
2012-2014
\$100,000

Canadian Institutes of Health Research Prize in Aging
Award
2012-2014
\$5000.00

Queen Elizabeth II Graduate Scholarship in Science and
Technology
2012-2013
\$15,000

Ontario Graduate Scholarship
2012-2013
(Declined) \$15,000

Graduate Student Award Parkinson Society Canada
2010-2012
\$40,000

Ontario Graduate Scholarship
2010-2011
(Declined) \$15,000

Western Graduate Research Scholarship 2009-2013
\$15,000/year
(4 years)

Research Excellence Award – Parkinson Society Southwest
Ontario
2008-2009.

Service Distinction Award – Arkansas Alzheimer's Society
2003-2004

Publication Activity During Doctoral Training

Manuscripts submitted or in preparation	8
Refereed Chapters in Books	2
Refereed Articles in Journals	12
Refereed Articles Accepted for publication (in press)	2
Refereed Articles in Conference Proceedings	2
Articles in non-Peer-Reviewed Journals, Conference Proceedings, or Government/Technical Writings	10
Published Refereed Abstracts	45

Published or Accepted Refereed Articles

1. South A., Somers S., Jog M. Gum chewing improves swallow frequency and latency in Parkinson patients: A preliminary study. *Neurology*. 2010; 74: 1198-1202.
 2. South, A. and Orange, JB. Word retrieval deficits in amyotrophic lateral sclerosis (ALS). In Clinical Aphasiology Conference Proceedings (2010; Isle of Palms May 23-27, 2010). Publisher: The Aphasiology Archive, University of Pittsburgh. <http://aphasiology.pitt.edu/archive/00002171/>
 3. Rahimi F., Duval C., Jog M., Bee C., South A., Jog M., Edwards R., Boissy P. Capturing whole-body mobility of patients with Parkinson disease using inertial motion sensors: expected challenges and rewards. 2011; *Conference Proceedings IEEE Engineering in Medicine and Biology Society*; 5833-5837.
-

4. Kumar H., South A., Jog M. Carbamazepine responsive dysphagia- a variant of acquired Paroxysmal Kinesigenic Dyskinesia? *Movement Disorders*. 2011; 26(7): 1356-1358.
 5. Roberts-South A., Findlater K., Strong M., Orange JB. Longitudinal changes in discourse production in amyotrophic lateral sclerosis. *Seminars in Speech and Language*. 2012; 33(1): 79-94.
 6. South A., Findlater, K., Strong, M., Orange, JB. Longitudinal and discourse changes in amyotrophic lateral sclerosis. 2012. *Procedia Social and Behavioural Sciences*, volume 31.
 7. Rigby, H., Roberts-South, A., Kumar, H., Cortese, L., Jog, M. Diagnostic challenges revealed from a neuropsychiatry movement disorders clinic. *The Canadian Journal of Neurological Sciences*. 2012; 39(6): 782-788.
 8. South A., Hall L., Jog M. Telehealth Management of Movement Disorders Patients: A Retrospective Study. *The Canadian Journal of Neurological Sciences*. 2013; 40(2): 230-234.
 9. Rahimi F., Bee C., South A., Debicki D., Jog M. Variability of hand tremor in rest and in posture – a pilot study. *Conference Proceedings IEEE Engineering in Medicine and Biology Society*; 2013; 470-473.
 10. Rahimi, F., Bee, C., Debecki, D., Roberts, A., Bapat, P., Jog, M. Effectiveness of BoNT A in Parkinson's disease upper limb tremor management. *The Canadian Journal of Neurological Sciences*. 2013; 40(5): 663-669.
 11. Hossain I., Roberts-South, A., Jog, M., El-Sakka, M. Semi-automatic assessment of hyoid bone motion in digital videofluoroscopic images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*. 2013: 1-13.
 12. Bell Boucher, D., Roberts-South, A., Garcia, A. A., Katchabaw, M., & Jog, M. S. 2013; November). Immersive augmented reality: Investigating a new tool for Parkinson disease rehabilitation. In *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on* (pp. 1570-1573). IEEE.
 13. Spratt, J., Roberts, A., Eaton, B., Medzon, D. and Orange, J.B. 2013. A multimodal neurolinguistic treatment approach for conduction aphasia: A single case study. In *Clinical Aphasiology Conference: Clinical Aphasiology Conference* (2013: 43rd : Tucson, AZ : May 28-June 2, 2013) / : (2013). The Aphasiology Archive, University of Pittsburgh. <http://aphasiology.pitt.edu/archive/00002434/>
 14. Garcia, A.A., Roberts-South, A., Bell Boucher, D., Andre, N., Jackman, M., Katchabaw, M., & Jog, M.S. Advancing Immersive Augmented Reality in Parkinson Disease. *Virtual Reality*. Accepted February 2014.
 15. South A., Findlater K., Strong M., Orange JB. Word retrieval deficits in amyotrophic lateral sclerosis. *Aphasiology*, in press.
-

Published Refereed Book Chapters

1. Roberts, A. and Orange, JB. Discourse in Lowy body spectrum disorder. (2013). In *Pragmatics in Dementia Discourse*, Eds. Boyd Davis and Jacqueline Guendouzi. (pp. 153-211). Cambridge Scholars Publishing: Newcastle upon Tyne. (Invited peer reviewed submission).
2. Garcia A., Andre N., Bell Boucher, D., Roberts-South, A., Jog, M., Katchabaw, M. (2014). Immersive Augmented Reality for Parkinson Disease Rehabilitation. In *Virtual and Augmented Reality in Healthcare*. Springer Editorial Board. Springer Book Publishing.

Peer Reviewed Manuscripts under review or in preparation

1. Roberts, A., and Orange, JB. Conversation Difficulties in Parkinson disease dementia: A family perspective.
2. Nguyen, P., Roberts, A., Orange, JB., Jog., M., McRae, K. Parkinson's Patients Upper and Lower Limb Motor Impairments Differentially Influence Action Verb Processing
3. Graham, R., Teasell, R., Roberts, A., Orange JB. Survey of Canadian Speech-Language Pathologist Practices in Stroke Recovery Part One: Language and Communication Screening and Assessment.
4. Graham, R., Teasell, R., Roberts, A., Orange JB. Survey of Canadian Speech-Language Pathologist Practices in Stroke Recovery Part Two: Language and Communication Interventions.
5. Roberts, A., Penner, J., Debicki, D., Gatti, J., Jog, M., Bartha, R. 7-Tesla MR Spectroscopy in Parkinson Disease.
6. Gofton, T., Roberts, A., Kumar, H., Speechley, M., Jog, M. Validity, reliability and insights from implementation of the McGill Quality of Life Questionnaire in patients living with Parkinson's disease (MQOL-PD). Under review *Journal of Palliative Medicine*.
7. Rahimi, F., Bee, C., Roberts, A., Debecki, D., Bapat, P., Jog, M. Dynamic Decomposition of Motion in Essential and Parkinsonian Tremor. Under review *Canadian Journal of Neurological Sciences*.
8. Roberts, A., Solomon, A., Bell-Boucher, D., Kumar, H., Jog, M. The Influence of Levodopa and Cognition on Turning Strategies in PD. Under review *Journal of Parkinson Disease*.

Relevant Work Experience

Course Lecturer/Limited Duties Appointments Western University

CSD 9617: Acquired Language Disorders (Winter 2012; Winter 2013; Winter 2014; Winter 2015 (renewed))

CSD 9627S – Clinical Applications Course for Acquired Language Disorders (Winter 2012; Winter 2013; Winter 2014; Winter 2015 (renewed))

CSD 9602L section 004: Clinical Practicum in Speech Pathology Fluency (Spring 2012; Spring 2013); CSD 9604: Clinical Practicum in Speech Pathology Adult Neurogenic Disorders Individual and Group (Fall 2013)

Graduate Teaching Assistant Western University

CSD 9622B: Dysarthria, Dyspraxia, and Associated Disorders (Winter, 2010; Winter, 2011)

CSD 9610: Acquired Language Disorders (Fall, 2009; Fall, 2010)

CSD 9627: Acquired Language Disorders (Winter 2012)

CSD 9617: Acquired Language Disorders Clinical Applications (Winter 2012)

Clinical Speech and Language Pathology Positions

London Health Sciences Centre, London

Speech Language Pathologist Clinical Neurosciences

November 2004– September 2009

Levi Hospital, Hot Springs Arkansas

Clinical Supervisor and Speech Language Pathologist

Clinical Neurology Inpatient & Outpatient services

September 2002-November 2004

Aegis Therapies, Hot Springs Arkansas

Speech Language Pathologist/Manager

Adult Neurodegenerative disorders

June 2001-August 2002

Diagnostic Swallowing Services

Director of Clinical Services

May 2000-May 2001

Triumph Therapy Services

Owner/Director of Clinical Services

Feb 2000-Feb 2001

Vencare Rehabilitation Services Rehabilitation

Program Manager Adult Neurological Disorders

February 1998-February 2000

Jackson Madison County General Hospital

Speech Language Pathologist Adult Neurology

May 1995 - December 1997

Referred Grants/Funding

2014-2016 Parkinson Society Canada Pilot Grant Program

PI: Dr. Ken McRae

Role: Co-investigator

Project: “She will drive the ____”: Predictive language comprehension in persons with Parkinson disease”

Agency: Parkinson Society Canada

Funding: \$40,915 **Awarded**

2013-2015 STRIVE-Home: Stroke Rehabilitation Involving Videconferencing Element at Home.

PI: Dr. Robert Teasell

Role: Co-investigator

Project: Evaluating the efficacy and cost effectiveness of delivering speech pathology services post-stroke via telehealth

Agency: Heart and Stroke Foundation

Funding 1 year \$100,000 **Awarded**

2013-2018 Ontario Brain Institutes Neurodegenerative Disease Project (ONDRI)

PI: Dr. Michael Strong and the OBI Steering Committee

Role: Member of the Neuropsychology, Language, & Speech Platform.

Project: Longitudinal project evaluating genetic, neuroimaging, and phenotypic presentations across neurodegenerative disorders with cognitive decline and dementia: Parkinson disease, Alzheimer’s disease, frontotemporal dementia, ALS, vascular cognitive impairment.

Funding 5 years \$28.5 million **Awarded**

2012-2014 Canadian Institutes of Health Research Fellowship Research Award & Prize in Aging

Role: Co-investigator with JB Orange

Project Title: Non-Motor Communication Impairments in Parkinson’s Disease and The Effect On Quality Of Life For People With PD and Their Family Caregivers

Year 1: \$15,000 **Awarded**

Year 2: \$5,000 **Awarded**

2009 Parkinson’s Society Canada Pilot Grant Program

Role: Co-investigator with Dr. Mandar Jog

Project Title: Sensorimotor Priming in Swallowing in Parkinson’s disease

\$32,500 **Awarded**
